



UNIVERSITÉ DE
SHERBROOKE

Faculté de génie

Département de Génie Civil

ESTIMATION DE L'HUMIDITÉ DU SOL À L'AIDE D'IMAGES RADARSAT-2 ET DE
RÉSEAUX DE NEURONES : APPLICATION AUX BASSINS VERSANTS TRENT ET
SEVERN, ONTARIO

Mémoire de maîtrise

Spécialité: génie civil

Guillaume DESBIENS

Jury:

Robert Leconte (directeur)

François Cavayas, Université de Montréal

Kalifa Goïta, Université de Sherbrooke

RÉSUMÉ

L'humidité du sol joue un rôle important dans le partitionnement de l'eau entre l'infiltration et le ruissellement de surface, qui influence directement les débits en rivière et les niveaux des réservoirs. La connaissance de la distribution spatiale de l'humidité du sol permet donc d'optimiser les différents usages de la ressource en eau en périodes sèches et d'aider la prévision et la gestion d'inondations lors de fortes pluies. La grande variabilité spatiale de l'humidité du sol rend toutefois difficile l'utilisation de capteurs *in situ* pour en faire le suivi sur de grands territoires tels que les bassins versants. La télédétection peut apporter une valeur ajoutée de par son potentiel pour estimer l'humidité du sol à l'échelle du bassin versant.

Le présent projet porte sur l'estimation de l'humidité du sol dans les bassins versants Trent et Severn en Ontario qui possèdent une superficie combinée de 18 360 km². L'approche adoptée est basée sur les réseaux de neurones artificiels (RNA). Deux approches ont été évaluées. La première, l'*approche polarisation simple et double* utilise uniquement des données radar acquises en polarisation HH ou HV. La seconde approche, soit l'*approche polarimétrique*, utilise des données en polarisation HH, HV et VV en plus de paramètres polarimétriques. Au total, 37 images RADARSAT-2 ont été acquises en différentes polarisations et résolutions spatiales entre les mois de mai 2012 et août 2013. En plus des coefficients de rétrodiffusion radar, des données de pente, de texture du sol et de végétation ainsi que des paramètres obtenus suite à une décomposition polarimétrique de la cible ont été utilisés comme intrants aux RNA. Des cartes d'humidité du sol moyenne et d'incertitude, représentant, dans l'ordre, la moyenne et l'écart-type des estimations faites par les 30 RNA sélectionnés, ont été produites. Les performances et les cartes obtenues ont été analysées afin de déterminer l'approche la plus avantageuse pour cartographier l'humidité du sol à l'échelle du bassin versant.

Ce projet de recherche a illustré le potentiel, mais aussi les enjeux, liés à l'estimation de l'humidité du sol à l'échelle du bassin. Il a été démontré que, dans un contexte opérationnel, l'*approche polarisation simple et double* est la plus avantageuse. Les cartes d'humidité produites avec l'*approche polarimétrique*, plus coûteuses, n'ont pas montré d'améliorations statistiquement significatives par rapport à l'*approche polarisation simple et double*. De tous les RNA testés, celui offrant la meilleure performance utilise l'angle d'incidence et les coefficients de rétrodiffusion radar HH et HV comme données d'entraînement. Il a aussi été démontré que l'incertitude sur l'estimation de l'humidité du sol est étroitement liée aux données d'entraînement. Le recours à des variables statiques dans le temps, comme la texture du sol, a affecté négativement et de manière importante les cartes d'humidité du sol, et ce, malgré de bonnes performances selon des critères statistiques comme le coefficient de Pearson et l'erreur quadratique moyenne. L'analyse visuelle des cartes d'humidité du sol demeure donc un moyen privilégié pour évaluer la performance des RNA.

Des pistes de recherche sont suggérées en vue d'améliorer la performance des RNA. Une première avenue serait le recours à une plus grande quantité de données pour leur entraînement, qui pourraient être générées à partir de modèles de rétrodiffusion à base physique. On pourrait également corriger les coefficients de rétrodiffusion pour diminuer l'influence de la végétation sur le signal rétrodiffusé avant l'entraînement des réseaux.

Mots-clés : Télédétection, radar, humidité du sol, hydrologie, réseau de neurones

REMERCIEMENTS

La réalisation de ce projet est le fruit de la collaboration de nombreuses personnes et n'aurait pas pu être menée à terme sans eux et je tiens à souligner leur contribution.

Je veux tout d'abord remercier mon directeur de recherche, M. Robert Leconte, pour m'avoir offert l'opportunité de travailler sur ce projet de recherche aussi passionnant que complexe ainsi que pour ses conseils et son soutien. Pour ses conseils et son expertise, je tiens aussi à remercier Mme Mélanie Trudel.

J'aimerais aussi remercier les responsables des différentes *Conservation Authorities* ontariennes pour leur collaboration sans laquelle il aurait été impossible de collecter autant de données sur un si grand territoire. Aussi, j'aimerais remercier le ministère des Ressources Naturelles de l'Ontario pour avoir généreusement mis à ma disposition une mosaïque d'ortho photographies aériennes captées en 2008 couvrant la majorité des bassins versants Trent et Severn.

Le présent projet de recherche a été réalisé dans le cadre du Programme de bourses de recherche en milieu de pratique (BMP-Innovation) financé par le Fonds québécois de la recherche sur la nature et les technologies (FQRNT) et le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG). Je tiens aussi à remercier la firme de consultants AECOM, plus particulièrement MM. Denis Baron et François Trudeau, pour leur soutien financier, leur soutien technique et pour m'avoir offert un poste de travail au sein de leur équipe.

L'Agence spatiale canadienne est aussi impliquée puisque le projet global, qui inclut mon projet, est exécuté dans le cadre du Programme de développement d'applications en observation de la Terre (PDAOT). À terme, Parcs Canada sera l'utilisateur final de l'outil développé.

Pour terminer, je tiens à remercier ma conjointe Julie pour sa patience et son support moral qui m'ont permis de passer à travers les périodes plus difficiles et les passages à vide de motivation.

TABLE DES MATIÈRES

RÉSUMÉ	i
REMERCIEMENTS	iii
TABLE DES MATIÈRES.....	v
LISTE DES FIGURES	ix
LISTE DES TABLEAUX	xiii
CHAPITRE 1 INTRODUCTION.....	1
1.1 Mise en contexte et problématique	1
1.2 Définition du projet de recherche.....	3
1.3 Objectifs du projet de recherche	4
1.4 Contributions originales.....	4
1.5 Plan du document.....	4
CHAPITRE 2 ÉTAT DE L'ART.....	7
2.1 Humidité du sol.....	7
2.2 Télédétection radar.....	9
2.2.1 Aspects théoriques.....	9
2.2.2 Télédétection radar de l'humidité du sol	12
2.3 Méthodes d'estimation de l'humidité du sol.....	21
2.3.1 Inversion de modèle de rétrodiffusion.....	21
2.3.2 Réseau de neurones artificiels	24
2.3.3 RNA et estimation de l'humidité du sol	25
CHAPITRE 3 SITE D'ÉTUDE ET DONNÉES.....	31
3.1 Zone d'étude	31
3.2 Données sur les bassins versants Trent et Severn	34
3.2.1 Données de texture du sol.....	34

3.2.2	Données de pente.....	34
3.3	Données <i>in situ</i>	35
3.3.1	Données d'humidité du sol.....	35
3.3.2	Données météorologiques	37
3.4	Données de télédétection	37
3.4.1	Données optiques	37
3.4.2	Données radar à synthèse d'ouverture (RSO)	39
CHAPITRE 4	MÉTHODOLOGIE.....	43
4.1	Traitement des images RSO	45
4.1.1	Création des images d'angle d'incidence.....	45
4.1.2	Étalonnage radiométrique des images RSO	45
4.1.3	Décomposition polarimétrique	46
4.1.4	Orthorectification et correction radiométrique.....	47
4.2	Traitement des données auxiliaires.....	47
4.3	Création des polygones associés aux sondes	48
4.4	Création des intrants des réseaux de neurones.....	49
4.5	Développement des réseaux de neurones artificiels (RNA)	51
4.5.1	Sélection des variables entrantes aux RNA.....	51
4.5.2	Architecture des réseaux	53
4.5.3	Entraînement, validation et test des RNA	55
4.5.4	élection du critère de performance	59
4.5.5	Sélection des RNA	59
4.6	Création des cartes d'humidité du sol estimée et d'incertitude	62
CHAPITRE 5	RÉSULTATS ET DISCUSSION.....	67
5.1	Approche polarisation simple et double	67

5.1.1	Angle d'incidence et coefficient de rétrodiffusion HH	67
5.1.2	Angle d'incidence, coefficient de rétrodiffusion HH et pente.....	80
5.1.3	Angle d'incidence, coefficient de rétrodiffusion HH et texture	84
5.1.4	Angle d'incidence, coefficients de rétrodiffusion HH et HV	88
5.2	Approche polarimétrique	93
5.2.1	Angle d'incidence et puissance de diffusion de surface.....	94
5.2.2	Angle d'incidence, puissance de diffusion de surface et de volume	101
5.2.3	Angle d'incidence, coefficient de rétrodiffusion de surface HH.....	110
5.3	Influence de la taille de l'échantillon	118
5.3.1	Cartes d'estimation de l'humidité du sol.....	120
CHAPITRE 6 CONCLUSION		125
LISTE DES RÉFÉRENCES.....		129
ANNEXE A. CARTOGRAPHIE DE LA ZONE D'ÉTUDE		141
A.1	ÉPAISSEUR DU DÉPÔT MEUBLE.....	143
A.2	MATÉRIEL DE SURFACE DOMINANT	145
A.3	PERMÉABILITÉ DU SOL	147
A.4	ÉPAISSEUR DU DÉPÔT DE SURFACE.....	149
A.5	ORDRE ET GRANDS GROUPES DE SOLS	151
A.6	OCCUPATION DU SOL	153
A.7	CLIMAT	155
A.8	TEXTURE DU SOL.....	157
ANNEXE B. HISTOGRAMME DE FRÉQUENCE DE PENTE		159
ANNEXE C. SPÉCIFICATIONS DES MNÉ		161

LISTE DES FIGURES

Figure 2.1 Les différents mécanismes de diffusion, modifiée de Kornelson et Coulibaly [2013]	10
Figure 2.2 Contribution de la rétrodiffusion d'un couvert végétal sur un sol nu. Adaptée de Henderson et Lewis [1998, p. 427]	17
Figure 3.1 Localisation de la zone d'étude	31
Figure 3.3 a) Diagramme de la sonde EC-5 (adapté de Decagon Devices, Inc. [2001]) et b) station de mesure	36
Figure 3.4 Dispersions des données d'humidité du sol observée	37
Figure 3.5 Empreintes au sol des images RADARSAT-2 acquises	39
Figure 4.1 Organigramme méthodologique global	43
Figure 4.2 Organigramme de traitement des images RADARSAT-2	44
Figure 4.3 Organigramme du développement des RNA	44
Figure 4.4 Histogrammes de fréquence de σ° (a), de puissance (b) et de σ° de surface (c)	50
Figure 4.5 Histogrammes de fréquence de texture (a), d'humidité estimée, d'angle d'incidence et de pente (b), d'IVR et ISF (c)	51
Figure 4.6 Structure des réseaux de neurones développés	54
Figure 4.7 Humidité du sol et précipitations mesurées in situ en 2012 (a) et 2013 (b)	63
Figure 4.8 Carte de localisation des images utilisées pour la production des cartes d'estimation d'humidité ainsi que des sondes et stations météorologiques	65
Figure 4.9 Carte de l'utilisation des sols dans la zone de recouvrement des images RADARSAT-2 (source: [Ontario Ministry of Natural Resources, 2008])	66
Figure 5.1 Régression linéaire du meilleur RNA avec comme intrants l'angle d'incidence et σ°_{HH}	68
Figure 5.2 Résidus en fonction de l'humidité du sol (haut) et distribution des observations (bas) pour le $RNA_{\theta^\circ, \sigma^\circ_{HH}}$	69

Figure 5.3 Carte d'estimation de l'humidité du sol avec $RNA_{\theta^\circ, \sigma^\circ HH}$ à faibles angles d'incidence et superposition des sols très mal drainés pour le 1 ^{er} mai 2012 – conditions humides (a1 et a2), le 25 mai 2012 – conditions sèches (b1 et b2) et le 31 juillet 2013 – conditions sèches (c1 et c2).....	71
Figure 5.4 Carte d'estimation de l'humidité du sol avec $RNA_{\theta^\circ, \sigma^\circ HH}$ et superposition des sols très mal drainés à angles d'incidence élevés pour le 15 juillet 2012 – conditions sèches (a1 et a2) et le 8 août 2012 – conditions humides (b1 et b2)	72
Figure 5.5 Histogrammes de l'humidité estimée avec $RNA_{\theta^\circ, \sigma^\circ HH}$ pour les sous-bassins versants Baldwin et Pefferlaw	73
Figure 5.6 Cartes d'incertitude sur l'estimation de l'humidité du sol obtenues avec $RNA_{\theta^\circ, \sigma^\circ HH}$	75
Figure 5.7 Incertitude en fonction de $\sigma^\circ HH$	77
Figure 5.8 Histogrammes d'incertitude pour les sous-bassins versants Baldwin et Pefferlaw pour les $RNA_{\theta^\circ, \sigma^\circ HH}$	79
Figure 5.9 Régression linéaire du $RNA_{\theta^\circ, \sigma^\circ HH, pente}$	81
Figure 5.10 Carte d'estimation de l'humidité du sol avec avec $RNA_{\theta^\circ, \sigma^\circ HH, pente}$	82
Figure 5.11 Agrandissement d'une zone de la carte d'humidité estimée (a) et d'incertitude (b) obtenues pour le 15 juillet 2012	83
Figure 5.12 Histogrammes d'incertitude pour les sous-bassins versants Baldwin et Pefferlaw avec $RNA_{\theta^\circ, \sigma^\circ HH, pente}$	83
Figure 5.13 Régression linéaire du meilleur $RNA_{\theta^\circ, \sigma^\circ HH, texture}$	85
Figure 5.14 Cartes des fractions de sable (a) et d'argile (b).....	86
Figure 5.15 Carte d'estimation de l'humidité du sol $RNA_{\theta^\circ, \sigma^\circ HH, \% \text{ sable}, \% \text{ argile}}$	87
Figure 5.16 Régression linéaire du meilleur $RNA_{\theta^\circ, \sigma^\circ HH}$ et $\sigma^\circ HV$	88
Figure 5.17 Carte d'estimation de l'humidité du sol avec $RNA_{\theta^\circ, \sigma^\circ HH, \sigma^\circ HV}$	90
Figure 5.18 Histogrammes d'humidité estimée avec les $RNA_{\theta^\circ, \sigma^\circ HH}$ et $\sigma^\circ HV$ pour les sous-bassins versants Baldwin et Pefferlaw	91

Figure 5.19 Carte d'incertitude sur l'estimation de l'humidité du sol obtenues avec $RNA_{\theta^{\circ}, \sigma^{\circ}HH}$ et $\sigma^{\circ}HV$	93
Figure 5.20 Régression linéaire du $RNA_{\theta^{\circ}, Ps}$	95
Figure 5.21 Résidus en fonction de l'humidité du sol (haut) et distribution des observations (bas)) pour le $RNA_{\theta^{\circ}, Ps}$	95
Figure 5.22 Carte d'estimation de l'humidité du sol avec $RNA_{\theta^{\circ}, Ps}$	97
Figure 5.23 Variations d'humidité du sol estimée entre les $RNA_{\theta^{\circ}, \sigma^{\circ}HH}$, $\sigma^{\circ}HV$ et $RNA_{\theta^{\circ}, Ps}$	98
Figure 5.24 Histogrammes de l'humidité estimée avec $RNA_{\theta^{\circ}, Ps}$ pour les sous-bassins versants Baldwin et Pefferlaw	100
Figure 5.25 Régression linéaire du $RNA_{\theta^{\circ}, Ps, Pv}$	102
Figure 5.26 Carte d'estimation de l'humidité du sol avec $RNA_{\theta^{\circ}, Ps, Pv}$	104
Figure 5.27 Histogrammes de l'humidité estimée avec $RNA_{\theta^{\circ}, Ps, Pv}$ pour les sous-bassins versants Baldwin et Pefferlaw	106
Figure 5.28 Images de décomposition polarimétrique composé RVB – Pd; Pv; Ps	108
Figure 5.29 Carte d'incertitude sur l'estimation de l'humidité du sol obtenues avec $RNA_{\theta^{\circ}, Ps, Pv}$	110
Figure 5.30 Régression linéaire du meilleur $RNA_{\theta^{\circ}, \sigma^{\circ}Surf_HH}$	112
Figure 5.31 Humidité du sol observée par rapport à $\sigma^{\circ}Surf_HH$ en fonction de l'angle d'incidence	112
Figure 5.32 Humidité du sol observée par rapport à $\sigma^{\circ}HH$ en fonction de l'angle d'incidence	113
Figure 5.33 Carte d'estimation de l'humidité du sol avec $RNA_{\theta^{\circ}, \sigma^{\circ}Surf_HH}$	115
Figure 5.34 Histogrammes de l'humidité estimée avec $RNA_{\theta^{\circ}, \sigma^{\circ}Surf_HH}$ pour les sous-bassins versants Baldwin et Pefferlaw	116
Figure 5.35 Histogrammes d'incertitude pour les sous-bassins versants Baldwin et Pefferlaw avec $RNA_{\theta^{\circ}, \sigma^{\circ}Surf_HH}$	118

Figure 5.36 Régressions linéaires des meilleurs $RNA_{\theta^\circ, \sigma^\circ HH}$ utilisant 119 (gauche), 208 (centre) et 534 (droite) données	119
Figure 5.37 Histogrammes de distribution des valeurs d'angles d'incidence	120
Figure 5.38 Carte d'estimation de l'humidité du sol $RNA_{\theta^\circ, \sigma^\circ HH}$ en fonction de la taille de l'échantillon	123

LISTE DES TABLEAUX

Tableau 3.1 Images d'ISF acquises	38
Tableau 3.2 Sommaire des caractéristiques des images RADARSAT-2 acquises.....	40
Tableau 4.1 Sommaire des paramètres utilisés pour l'entraînement, la validation et le test.....	59
Tableau 4.2 Conditions observées lors de l'acquisition des images RADARSAT-2.....	64
Tableau 5.1 Sommaire des statistiques des 30 meilleurs $RNA_{\theta^{\circ}, \sigma^{\circ}HH}$	69
Tableau 5.2 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double.....	80
Tableau 5.3 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double.....	84
Tableau 5.4 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double.....	88
Tableau 5.5 Sommaire des statistiques des 30 meilleurs RNA pour l'approche de polarisation simple et double.....	89
Tableau 5.6 sommaire des variations entre les $RNA_{\theta^{\circ}, \sigma^{\circ}HH}$ et $RNA_{\theta^{\circ}, \sigma^{\circ}HH, \sigma^{\circ}HV}$	91
Tableau 5.7 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique	94
Tableau 5.8 Sommaire des statistiques des 30 meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique	96
Tableau 5.9 Sommaire des probabilités à angles d'incidence faibles et en présence de végétation abondante pour les $RNA_{\theta^{\circ}Ps}$	99
Tableau 5.10 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique	102
Tableau 5.11 Sommaire des statistiques des 30 meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique	103

Tableau 5.12 Sommaire des probabilités à angles d'incidence élevés et en présence de végétation abondante pour les $RNA_{\theta^{\circ}Ps,Pv}$	107
Tableau 5.13 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique.....	111
Tableau 5.14 Sommaire des statistiques des 30 meilleurs RNA RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique.....	113
Tableau 5.15 Sommaire des probabilités à angles d'incidence faibles et en présence de végétation abondante pour les $RNA_{\theta^{\circ}Ps,Pv}$	114
Tableau 5.16 Performances du meilleur RNA en fonction de la taille de l'échantillon.....	119
Tableau 5.17 Sommaire des statistiques des 30 meilleurs RNA en fonction de la taille de l'échantillon.....	121

CHAPITRE 1 INTRODUCTION

1.1 Mise en contexte et problématique

Les bassins versants des rivières Trent et Severn occupent une superficie de 18 360 km² entre la baie Georgienne et le lac Ontario dans la province canadienne de l'Ontario. Ces deux bassins versants sont traversés par la voie navigable Trent-Severn sur une longueur de 386 kilomètres au long de laquelle se trouvent 160 barrages et ouvrages de contrôle des débits. La voie navigable étant la propriété de Parcs Canada, l'opération des ouvrages de régulation est sous sa responsabilité. L'eau qui se trouve sur son territoire est utilisée pour de nombreux usages; elle sert d'approvisionnement en eau aux centrales hydroélectriques, des municipalités y pompent leur eau de consommation, les agriculteurs l'utilisent pour leurs activités, en plus des milliers de plaisanciers qui l'empruntent annuellement.

Au cours des dernières années, la concurrence pour l'usage de l'eau a augmenté de façon importante sur le territoire drainé par les bassins versants Trent et Severn. Afin de faire face à cette problématique, Parcs Canada désire mettre en œuvre une vision globale de son programme d'amélioration de la gestion de l'eau tout en s'assurant que ses employés aient les outils nécessaires pouvant les aider dans leurs prises de décision relativement à la gestion de l'eau. Dans cette optique, la firme AECOM a été mandatée par Parcs Canada pour la réalisation d'une étude de gestion de l'eau en 2010. Cette étude a mené à la proposition d'une approche relative à la gestion future de l'eau. Parmi les différents outils proposés qui permettraient une meilleure gestion de l'eau, Parcs Canada a priorisé la mise en place d'un outil de prévision hydrologique. Cet outil permettrait d'estimer les apports en eau à venir sur une plage temporelle de quelques jours.

Lors de précipitations, la quantité d'eau présente dans le sol contrôle le taux d'infiltration et, par conséquent, l'importance du ruissellement de surface subséquent qui influence directement les débits en rivières et les niveaux des réservoirs. Une bonne connaissance des conditions d'humidité du sol et de sa distribution spatiale a donc une importance majeure dans la gestion par bassins versants puisqu'elle permet d'optimiser l'allocation de la ressource en périodes sèches et d'aider la prévision et la gestion d'inondations lors d'événements pluviométriques extrêmes. Toutefois, l'humidité du sol est une variable très peu mesurée en grande partie parce

que sa grande variabilité spatiale et temporelle nécessiterait l'installation de nombreux capteurs pour permettre de comprendre la dynamique de cette variable d'état du bassin versant. Pour pallier au manque de données sur l'humidité du sol, des méthodes alternatives ont été développées pour estimer l'état hydrique du bassin versant, notamment pour fins de prévision hydrologique, comme les indicateurs de précipitation antécédente [Ali *et al.*, 2010; Beck *et al.*, 2009; Heggen, 2001], l'indice d'humidité du sol («*Soil Wetness Index*») [Brocca *et al.*, 2010] et l'indice topographique d'humidité («*Topographic Wetness Index*») [Grabs *et al.*, 2009; Sørensen *et al.*, 2006]. Aujourd'hui, en raison de la pression anthropique (utilisation du territoire, changements climatiques, etc.) sur la ressource en eau, la connaissance quantitative de l'humidité du sol et de son comportement spatial et temporel revêt une importance accrue.

Afin de mieux prendre en considération l'influence de l'humidité du sol et d'améliorer les prévisions hydrologiques, une approche basée sur l'insertion directe de données d'humidité du sol dans un modèle hydrologique a été développée dans le cadre du mandat confié à AECOM. Ces données d'humidité du sol sont estimées à partir de données acquises par le satellite RADARSAT-2 et de réseaux de neurones artificiels (RNA). Cette approche offre le potentiel d'obtenir des estimations de l'humidité du sol dans les premiers centimètres de surface pour une portion appréciable du territoire, contrairement à quelques données ponctuelles provenant de stations hydrométéorologiques. En contrepartie, la couverture temporelle est limitée par le temps de passage du satellite au-dessus de la région ciblée, qui peut être de plusieurs jours. Le présent projet de recherche porte sur l'estimation de l'humidité du sol en surface. L'intégration des estimations au modèle hydrologique ne sera donc pas abordée.

Avec la mise en orbite des premiers satellites d'observation de la Terre dans les années 1970, des mesures de surface des bassins versants, comme des indices de végétation (indice de surface foliaire, indice de végétation normalisé - NDVI, etc.) et l'étendue de la couverture de neige, sont de plus en plus disponibles pour aider à la gestion de la ressource en eau à l'échelle du bassin. Plus récemment, des satellites dotés de capteurs radar polarimétriques ont été mis en orbite, ce qui signifie qu'ils ont la capacité d'acquérir des données pour l'ensemble des polarisations (HH, VV, HV et VH) de même que la phase associée au signal. Ces capteurs, comme le satellite canadien RADARSAT-2, offrent un potentiel intéressant pour l'estimation de l'humidité du sol dans les premiers centimètres de sol à l'échelle du bassin versant, comme en font foi de nombreuses études à ce sujet [Altese *et al.*, 1996; Baghdadi *et al.*, 2012a; Barrett

et al., 2012; Bourgeau-Chavez *et al.*, 2013; Hajnsek *et al.*, 2003, 2009; Kornelsen et Coulibaly, 2013; Lakhankar *et al.*, 2006; McNairn *et al.*, 2010; Trudel *et al.*, 2012]. Plusieurs approches ont été développées pour convertir le signal radar rétrodiffusé en humidité du sol (voir l'état de l'art à la section 2.3.1), parmi lesquelles on retrouve des modèles empiriques relativement simples à implanter, mais dont le champ d'application est limité aux conditions pour lesquelles le modèle a été produit, ainsi que des modèles théoriques, plus universels mais dont l'implantation est plus compliquée de par la nécessité d'avoir beaucoup de données pour les faire fonctionner.

1.2 Définition du projet de recherche

Le présent projet de recherche porte sur l'estimation de l'humidité du sol à partir de données radar satellitaires. Il vise à fournir une évaluation quantitative de l'état hydrique des bassins versants Trent et Severn de manière à pouvoir par la suite effectuer une prévision hydrologique à court terme du bassin. Dans ce contexte spécifique, le recours à un modèle théorique d'obtention de l'humidité du sol à partir de données radar a été écarté pour favoriser une approche plus simple, en terme d'acquisition de données in situ et de calibration, qui se prête davantage à une utilisation opérationnelle de ces données. L'approche de modélisation retenue, basée sur les réseaux de neurones artificiels (RNA), a produit des résultats encourageants dans différentes études utilisant des données radar combinées à des données auxiliaires comme intrants à des RNA pour estimer l'humidité du sol [Baghdadi *et al.*, 2002a; Baghdadi *et al.*, 2016; Del Frate *et al.*, 2003; Notarnicola *et al.*, 2008; Said *et al.*, 2008]. Enfin, le recours à des données d'observation de la Terre est motivé dans ce projet par le vaste territoire faisant l'objet d'une gestion de la ressource en eau par Parcs Canada, l'accès parfois restreint au territoire ainsi que l'absence de mesures de terrain, outre des stations météorologiques, permettant de définir l'état hydrique du bassin versant.

Le projet de recherche vise donc à vérifier l'hypothèse suivante : est-ce que l'ajout de paramètres polarimétriques comme intrants aux RNA apportera une augmentation significative de la précision et une réduction de l'erreur sur l'estimation de l'humidité du sol?

1.3 Objectifs du projet de recherche

L'objectif principal de cette recherche est d'utiliser des données RADARSAT-2 afin d'estimer l'humidité du sol en surface pour des sols agricoles nus ou faiblement végétalisés à l'aide d'un modèle de type RNA.

Les objectifs spécifiques sont 1) d'évaluer les résultats obtenus selon deux approches, présentées ci-dessous, et de déterminer la plus avantageuse, 2) de quantifier l'incertitude sur les estimations d'humidité du sol et 3) de produire des cartes d'humidité du sol et d'incertitude de l'estimation.

Les deux approches évaluées sont 1) l'approche *Polarisation simple et double*, réalisée à l'aide de données co-polarisées horizontalement (HH) et en polarisation croisée (HV) et 2) l'approche *Polarimétrique*, à l'aide de données Quad-pol standard (HH, VV, VH et HV).

1.4 Contributions originales

Plusieurs recherches ont été réalisées dans le but d'estimer l'humidité du sol à l'aide de RNA utilisant des données radar et des variables auxiliaires comme intrants. Toutefois, peu d'études ont porté sur l'effet de l'utilisation de paramètres polarimétriques tels que la puissance de diffusion de surface (P_S), de volume (P_V), de double réflexion (P_D) et le coefficient de rétrodiffusion de surface en polarisation HH ($\sigma^{\circ}_{\text{Surf_HH}}$) comme intrants aux RNA sur la précision des estimations. Les résultats obtenus dans le cadre de cette recherche pourraient ensuite être intégrés à un modèle hydrologique par assimilation de données. Compte tenu que ce processus nécessite de fournir des valeurs d'incertitude, celle-ci sera déterminée et représentée sous forme de cartes. Les contributions originales de cette étude sont donc:

1. L'utilisation de paramètres polarimétriques comme intrants aux réseaux de neurones.
2. La création de cartes d'incertitude sur les estimations d'humidité du sol.

1.5 Plan du document

Le présent mémoire de recherche contient 6 chapitres. La mise en contexte, la problématique, les objectifs de recherche et les contributions originales ont été présentés dans le présent chapitre. L'état de l'art est présenté au chapitre 2 en trois sections principales : la télédétection radar, l'estimation de l'humidité du sol et les réseaux de neurones artificiels. Le chapitre 3 présente la zone d'étude ainsi que les données utilisées dans le cadre de ce projet. Le

chapitre 4 présente la méthodologie développée pour cette étude. Les résultats et l'analyse de ceux-ci sont présentés au chapitre 5. Les conclusions tirées du processus de recherche et des résultats sont présentées au chapitre 6.

CHAPITRE 2 ÉTAT DE L'ART

Le présent chapitre présente l'état actuel des recherches en lien avec les différents domaines impliqués dans ce projet de recherche. Dans un premier temps, un bref survol des recherches portant sur l'humidité du sol dans un contexte hydrologique est présenté. La deuxième partie porte sur la télédétection radar, les modèles de rétrodiffusion développés ainsi que les méthodes d'inversion d'humidité du sol utilisées. La troisième partie porte sur les réseaux de neurones artificiels et leur utilisation pour estimer l'humidité du sol.

2.1 Humidité du sol

Malgré le fait que l'humidité du sol en surface ne constitue que 0,0012% de l'eau disponible sur la Terre [Chow *et al.*, 1988], elle joue un rôle majeur dans le cycle hydrologique en affectant la partition des précipitations entre l'infiltration, le ruissellement de surface, la percolation et l'évapotranspiration. Ces dernières sont des variables d'état clé dans plusieurs applications hydrologiques telles que la modélisation et la prévision hydrologique, la prévision d'inondations ainsi que la gestion agricole et de la ressource hydrique. En effet, plusieurs études ont fait le lien entre les conditions antécédentes d'humidité du sol et le rapport ruissellement/précipitation [Cassardo *et al.*, 2002; Maurer *et al.*, 2004; Merz et Plate, 1997; Zehe et Blöschl, 2004]. Plus récemment, Thorstensen *et al.* [2016] ont intégré des données d'humidité du sol provenant d'un réseau de 20 sondes à la procédure de calage d'un modèle hydrologique distribué. Cette approche a permis de réduire l'erreur quadratique moyenne (RMSE) de 33% en moyenne sur l'humidité du sol estimée par le modèle, permettant du coup de mieux représenter l'hétérogénéité de cette variable à l'échelle du bassin versant. Par contre, le RMSE sur les débits simulés et le critère de Nash-Sutcliffe obtenus ne permettent pas d'affirmer que cette méthode de calage améliore la prévision des débits, suggérant ainsi la nécessité de recherches plus approfondies.

La variabilité spatiale et temporelle de l'humidité du sol fait en sorte que son suivi *in situ* est une approche nécessitant d'importantes ressources financières. La grande couverture au sol et la résolution spatiale que la télédétection offre permet de faire face à ces contraintes [Dubois *et al.*, 1995a; Lakhankar *et al.*, 2006; Paloscia *et al.*, 2013; Said *et al.*, 2008]. Les capteurs spatioportés ont la capacité de fournir une haute résolution spatiale et/ou temporelle et sont par

conséquent bien adaptés pour des études à l'échelle du bassin versant. Toutefois, les capteurs opérant dans la portion visible du spectre électromagnétique ne peuvent générer de l'information utilisable que dans des conditions de ciel dégagé et sous un éclairage solaire. De plus, l'information captée par ces capteurs provient uniquement d'une très mince couche située à l'interface air-sol. Dans le but de remédier en partie à ces contraintes, les données radar, acronyme pour *Radio Detection And Ranging*, ont été utilisées dans de nombreuses applications.

En plus de la variabilité temporelle et spatiale, l'humidité du sol varie aussi en profondeur. On distingue fréquemment deux zones distinctes : l'humidité de surface (les premiers centimètres) et celle de la zone racinaire (jusqu'à 1,5 m) [Capehart et Carlson, 1997; Paris Anguela *et al.*, 2008; Wilson *et al.*, 2003]. L'humidité de surface présente une variabilité très importante alors que celle en profondeur est faible. Cet écart s'explique par la plus grande influence que les conditions atmosphériques telles que la pluie, le vent et la radiation solaire ont sur l'humidité en surface [Paris Anguela *et al.*, 2008]. L'humidité présente dans la zone racinaire contrôle la transpiration de la végétation et est typiquement celle à laquelle on s'intéresse en modélisation hydrologique [Wilson *et al.*, 2003]. Dans cette optique, Paris Anguela *et al.* [2008] ont étudié la qualité de trois produits d'humidité du sol pouvant être utilisés en assimilation de données. Ils ont notamment comparé l'humidité *in situ* (m_{vo}) mesurée avec une sonde de type *Time Domain Reflectometry* (TDR) à celle dérivée de données radar (m_{ve}) en bande C (ERS) ayant une résolution de 25 km. L'index d'eau du sol (*Soil Water Index*, SWI) a ensuite été dérivé de m_{ve} puis converti en m^3/m^3 afin de représenter l'humidité de la zone racinaire. Le SWI a été obtenu à l'aide du modèle hydrique à deux couches proposé par Wagner *et al.* [1999] et représente le contenu en eau du premier mètre de sol en unité relative variant entre le point de flétrissement et la capacité au champ. L'humidité de surface estimée et *in situ* sont corrélées avec un coefficient de corrélation (R^2) de 0,53 et une erreur quadratique moyenne (RMSE) de 0,05 m^3/m^3 a été obtenue, ce qui est considéré satisfaisant compte tenu de l'importante variabilité de m_v présente dans cette zone. Un R^2 de 0,66 et une RMSE de 0,06 ont pour leur part été obtenus pour l'humidité de la zone racinaire estimée. Ces résultats suggèrent donc que les données d'humidité du sol obtenues par télédétection radar peuvent, une fois traitées, être utilisées à des fins de modélisations et de prévisions hydrologiques.

2.2 Télédétection radar

2.2.1 Aspects théoriques

Les systèmes de télédétection radar sont des capteurs actifs émettant leur propre rayonnement électromagnétique afin d'illuminer la surface ou la cible pour ensuite mesurer la quantité de radiation retournée au capteur [Fung, 1994; Ulaby *et al.*, 1982]. Les capteurs radar opèrent dans la partie micro-ondes du spectre électromagnétique, qui s'étend entre 0,3 GHz et 300 GHz. Les fréquences les plus utilisées en télédétection sont les bandes L (1–2 GHz), C (4–8 GHz) et X (8–12,5 GHz), qui correspondent à des longueurs d'ondes (λ) entre 15–30 cm, 3,75–7,5 cm et 2,4–3,75 cm, respectivement [Centre canadien de télédétection, 1997; Ulaby *et al.*, 1981]. Aux longueurs d'ondes comprises entre 3,75 et 30 cm, soit aux fréquences inférieures à 8 GHz, plus grandes que les ondes visibles et infrarouges, le signal est insensible à la diffusion atmosphérique occasionnée par les nuages, la pluie, la bruine et la poussière [Fung, 1994; Ulaby *et al.*, 1982]. Toutefois, à des fréquences supérieures à 8 GHz (la bande X par exemple), la pluie peut diffuser le signal de manière plus ou moins importante en fonction de la taille des gouttes. Ces caractéristiques permettent l'acquisition d'images radar en bandes C et L de jour comme de nuit et dans presque toutes les conditions atmosphériques.

Après être entrée en contact avec la surface, l'énergie incidente peut 1) être absorbée (ou atténuée), 2) être diffusée, 3) être réfléchie ou 4) pénétrer le médium pour ensuite être diffusée ou absorbée à nouveau [Ulaby *et al.*, 1982]. Trois mécanismes de rétrodiffusion sont observables : les rétrodiffusions de surface et de volume ainsi que la double réflexion sol-tronc. La figure 2.1 illustre ces différents cas de figure.

Ainsi, la quantité d'énergie rétrodiffusée par une surface naturelle, exprimée par le coefficient de rétrodiffusion, est fonction de la configuration du capteur (angle d'incidence (θ), longueur d'onde (λ) et polarisation de l'onde (p)), des propriétés de la surface (propriétés diélectriques et rugosité de surface) et des caractéristiques de la végétation (propriétés diélectriques et structure physique) [Ulaby *et al.*, 1982]. Les nombreux paramètres contribuant aux variations du coefficient de rétrodiffusion rendent donc l'estimation de l'humidité du sol à l'aide de données radar complexe.

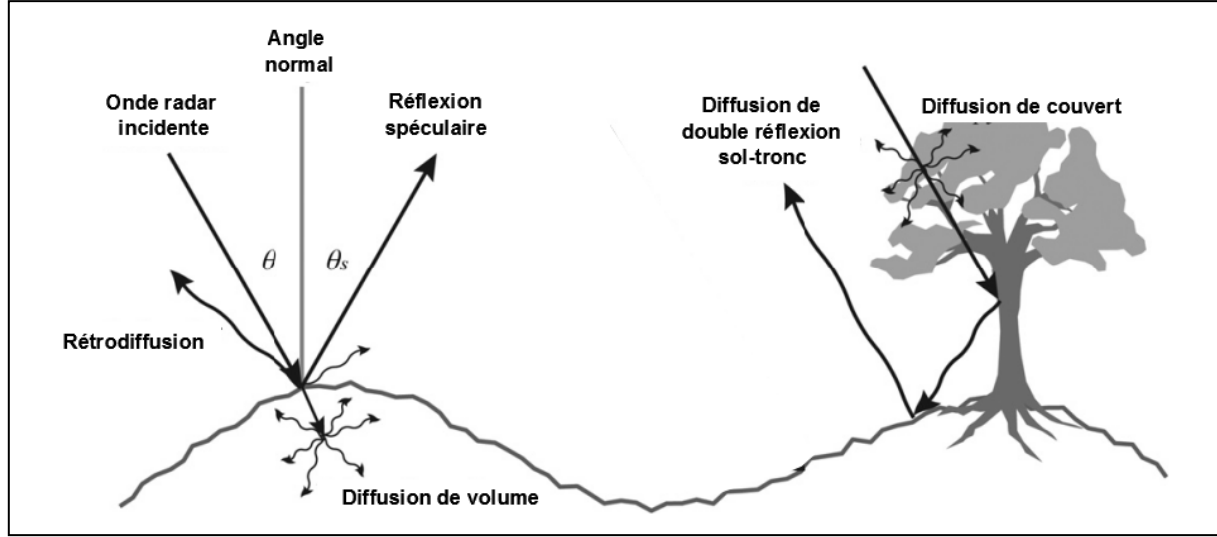


Figure 2.1 Les différents mécanismes de diffusion, modifiée de Kornelson et Coulibaly [2013]

Le signal rétrodiffusé peut être relié au signal incident à l'aide de la matrice de diffusion $[S]$ (équation (2.1)) :

$$\begin{bmatrix} E_H^d \\ E_V^d \end{bmatrix} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \begin{bmatrix} E_H^i \\ E_V^i \end{bmatrix} \quad (2.1)$$

où $[E_p^d]$ est le champ électrique de l'onde diffusée en polarisation p , et $[E_q^i]$ est le champ électrique de l'onde incidente dans la polarisation q .

Le coefficient de rétrodiffusion (σ_{qp}°) est ensuite relié à la matrice de diffusion à l'aide de l'équation (2.2) :

$$\sigma_{qp}^\circ = |S_{qp}|^2 \quad (2.2)$$

Les matrices de cohérence $[T]$ (équation (2.3)) et de covariance $[C]$ (équation (2.4)) permettent également de représenter la rétrodiffusion d'une cible et sont souvent utilisées par les méthodes de décomposition de la cible :

$$[T] = \frac{1}{2} \begin{bmatrix} |S_{HH}|^2 + 2\Re(S_{HH}S_{VV}^*) + |S_{VV}|^2 & |S_{HH}|^2 - 2j\Im(S_{HH}S_{VV}^*) - |S_{VV}|^2 & 2S_{HH}S_{HV}^* + 2S_{VV}S_{HV}^* \\ |S_{HH}|^2 + 2j\Im(S_{HH}S_{VV}^*) - |S_{VV}|^2 & |S_{HH}|^2 - 2\Re(S_{HH}S_{VV}^*) + |S_{VV}|^2 & 2S_{HH}S_{HV}^* - 2S_{VV}S_{HV}^* \\ 2S_{HV}S_{HH}^* + 2S_{HV}S_{VV}^* & 2S_{HV}S_{HH}^* - 2S_{HV}S_{VV}^* & 4|S_{HV}|^2 \end{bmatrix} \quad (2.3)$$

où \Re est la partie réelle et \Im est la partie imaginaire, et :

$$[C] = \frac{1}{2} \begin{bmatrix} |S_{HH}|^2 & \sqrt{2S_{HH}S_{HV}^*} & 2S_{HH}S_{VV}^* \\ \sqrt{2S_{HV}S_{HH}^*} & |S_{HV}|^2 & \sqrt{2S_{HV}S_{VV}^*} \\ S_{VV}S_{HH}^* & \sqrt{2S_{VV}S_{HV}^*} & |S_{VV}|^2 \end{bmatrix} \quad (2.4)$$

où $[E_p^d]$ est le champ électrique de l'onde diffusée en polarisation p, et $[E_q^i]$ est le champ électrique de l'onde incidente dans la polarisation q.

La polarisation se définit comme étant l'orientation du champ électrique d'une onde électromagnétique. Les antennes d'un système radar peuvent être configurées afin d'émettre et capter un rayonnement électromagnétique polarisé horizontalement ou verticalement. La plupart des satellites radar à synthèse d'ouverture (RSO) permettent l'acquisition de données en polarisations simple et double. La polarisation simple permet d'obtenir des données copolarisées (HH ou VV), et en polarisation croisée (HV ou VH). En double polarisation, les combinaisons suivantes sont possibles : HH + HV (i.e., 2 images, l'une en HH, l'autre en HV, peuvent être obtenues simultanément), VV + VH ou HH + VV.

Plus récemment, les capteurs RSO ALOS 2/PALSAR-2, TerraSAR-X et RADARSAT-2, opérant respectivement en bande L, X et C, ont été mis en orbite. Ces capteurs polarimétriques permettent l'acquisition de données en polarisation quadruple (HH, HV, VV et VH). Ils se distinguent des autres RSO de par leur capacité à mesurer la phase de l'onde en plus de l'amplitude, qui sont enregistrées en tant que nombres complexes. Il est donc possible de mesurer la différence de phase (ϕ) copolarisée ($\phi_{hh}-\phi_{vv}$) et croisée ($\phi_{hv}-\phi_{vh}$) [Ulaby *et al.*, 1996]. Ils permettent aussi d'obtenir les matrices de diffusion, de cohérence et de covariance de même que la signature polarimétrique pour l'ensemble des pixels d'une image, ce qui améliore la capacité à caractériser les propriétés physiques des objets et l'extraction de propriétés biophysiques et géophysiques de la surface terrestre [Lee et Pottier, 2009]. Plusieurs paramètres polarimétriques peuvent être obtenus à l'aide de technique de décomposition de la cible (*ex.* [Cloude et Pottier, 1997; Freeman et Durden, 1998; van Zyl, 1989; Yamaguchi *et al.*, 2005]). Les paramètres polarimétriques angle alpha (α) et anisotropie (A) de la décomposition Cloude–Pottier en bande C ont démontré leur capacité à discriminer les sols très humides ($m_v > 0,30 \text{ g/cm}^3$) des sols moins humides et les sols lisses des sols plus rugueux, respectivement [Baghdadi *et al.*, 2012a]. L'humidité du sol volumétrique $m_v \text{ (g/cm}^3\text{)}$ est présentée à l'équation 2.5.

$$m_v = \frac{W_h - W_d}{V} \quad (2.5)$$

où W_h est la masse humide (g), W_d la masse sèche (g) et V le volume (cm^3)

2.2.2 Télédétection radar de l'humidité du sol

Le sol est caractérisé par ses propriétés diélectriques ϵ et sa rugosité de surface. L'intérêt pour les capteurs radar dans les recherches sur l'estimation de l'humidité du sol est principalement dû à la sensibilité des micro-ondes aux propriétés diélectriques du sol, qui sont fonction du contenu en eau. En effet, la constante diélectrique (ϵ) d'un sol sec est de 3, indépendamment de la fréquence, alors que celle de l'eau est d'environ 80 [Dobson et Ulaby, 1986]. Pour des sols très humides, ϵ peut atteindre des valeurs supérieures à 20 [Schmugge, 1983]. C'est cet écart important qui rend possible l'estimation de l'humidité du sol à partir de données radar. En effet, Dobson et Ulaby [1981] ont observé une différence d'environ 10 dB entre un sol sec ($m_v = 0,139 \text{ g/cm}^3$) et un sol humide ($m_v = 0,456 \text{ g/cm}^3$) pour un angle d'incidence de 20° à une fréquence de 4,75 GHz (bande C). De plus, l'étude de Dobson et Ulaby [1981] a démontré que le coefficient de rétrodiffusion radar est proportionnel à m_v indépendamment de la fréquence et de l'angle d'incidence (voir la section 2.3.1).

Toutefois, lorsque le sol atteint le point de saturation en eau, soit environ $m_v = 0,35 \text{ g/cm}^3$, σ° perd cette sensibilité et demeure constant pour ensuite diminuer avec une augmentation de m_v . Cette diminution, liée à la plus grande réflexion spéculaire produite par les sols saturés et à l'indépendance de σ° à la rugosité de surface, a été observée dans plusieurs études (*ex.*: [Holah *et al.*, 2005; Bruckler *et al.*, 1988; Dobson et Ulaby, 1981]). Il est par conséquent difficile de cartographier avec précision l'humidité du sol à l'aide de données radar lorsqu'elle se situe au-delà de 0,35–0,40 g/cm^3 .

L'épaisseur de la couche du sol de laquelle provient la rétrodiffusion est égale à la profondeur de pénétration (δ_p), profondeur à laquelle la puissance de l'onde incidente est atténuée de $1/e$, ou environ 37 % [Ulaby *et al.*, 1982]. La profondeur de pénétration est fonction du contenu en eau du sol et de la longueur d'onde : elle augmente avec une diminution de m_v et de la fréquence [Ulaby *et al.*, 1996]. Il est généralement accepté qu'à des fréquences supérieures à 4 GHz (bandes K, X et C), un échantillonnage à une profondeur de 5 cm est optimal pour des

sols ayant une humidité volumétrique (m_v) supérieure à 5 %. La présence d'un gradient vertical d'humidité peut toutefois résulter en une surestimation de σ° [Boisvert *et al.*, 1995].

L'eau contenue dans le sol est soit adsorbée ou libre et c'est cette dernière qui influence la constante diélectrique de façon importante. La proportion de l'eau qui est adsorbée est proportionnelle à l'aire surfacique des particules composant le sol. Comme l'aire surfacique est inversement proportionnelle à la taille de la particule, les argiles possèdent la plus grande capacité d'adsorption. Ainsi, pour une même humidité, les sols sableux contiennent davantage d'eau libre et possèdent une constante diélectrique plus élevée que les sols argileux [Hallikainen *et al.*, 1985; Schmugge, 1983], ce comportement étant le plus prononcé à des fréquences inférieures à 5 GHz [Hallikainen *et al.*, 1985]. Toutefois, la constante diélectrique des sols secs est essentiellement indépendante de la texture et de la fréquence [Hallikainen *et al.*, 1985]. Bien que la constante diélectrique présente une sensibilité à la texture, celle-ci est relativement faible comparativement à m_v [Ulaby *et al.*, 1996].

Le comportement diélectrique des sols, et donc σ° , est influencé par la granulométrie puisque celle-ci détermine la quantité d'eau libre et que cette dernière influence la constante diélectrique de façon importante [Hallikainen *et al.*, 1985]. En effet, des sols plus sableux possèdent, pour une même humidité volumétrique, une plus grande proportion d'eau libre comparativement à ceux plus argileux, ce qui résulte en une constante diélectrique plus élevée et en une plus grande corrélation entre le coefficient de rétrodiffusion et l'humidité du sol pour les sols sableux [Blumberg *et al.*, 2000; Kong et Dorling, 2008; Srivastava *et al.*, 2006; Walker *et al.*, 2004]. Cela est dû au fait que dans les sols argileux, davantage de molécules d'eau sont adsorbées aux particules du sol, immobilisant les dipôles et les empêchant ainsi d'interagir avec le signal radar [Dobson et Ulaby, 1981]. La conversion de l'humidité du sol en eau libre a permis d'obtenir une corrélation de 0,96 avec l'estimation de l'humidité du sol, comparativement à 0,88 lorsque le contenu gravimétrique en eau a été considéré [Srivastava *et al.*, 2006].

Sols nus

Les premiers travaux portant sur l'estimation de l'humidité du sol à l'aide du signal radar ont été publiés en 1974 [Ulaby, 1974]. L'effet de l'humidité du sol sur le coefficient de rétrodiffusion (σ°) a été étudié en mesurant la réponse spectrale pour des fréquences variant de

4 à 8 GHz avec des angles d'incidence allant de 0° à 70° pour des sols agricoles nus légèrement rugueux et très rugueux, en terme de longueur d'onde. Les résultats obtenus indiquent que la réponse radar au contenu en eau du sol est fortement dépendante de la rugosité de surface, de la fréquence et de l'angle d'incidence (θ°).

Compte tenu de l'importance de la rugosité de surface sur σ° , de nombreuses recherches ont été réalisées afin d'en améliorer la modélisation et, à terme, améliorer les estimations d'humidité du sol. Lorsqu'une onde incidente frappe une surface avec un certain degré de rugosité, de l'énergie est diffusée dans toutes les directions, dont une partie vers le capteur, et le reste est réfléchi de manière spéculaire. Une augmentation de la rugosité entraîne une augmentation de la diffusion de surface ainsi qu'une diminution de la réflexion spéculaire, ce qui résulte en une augmentation du coefficient de rétrodiffusion [Ulaby *et al.*, 1982]. Plusieurs recherches réalisées en bande C sur des sols nus ont permis d'établir que la relation entre le coefficient de rétrodiffusion et la rugosité de surface suit une fonction exponentielle [Baghdadi *et al.*, 2008; Zribi et Dechambre, 2002]. De plus, cette relation tend à être plus forte à angles d'incidence élevés pour les rugosités de surface relativement faibles [Altese *et al.*, 1996; Baghdadi *et al.*, 2002b; Fung, 1994; Zribi et Dechambre, 2002].

La sensibilité du signal radar à la rugosité est aussi fonction de la longueur d'onde et de l'angle d'incidence local du signal [Fung *et al.*, 1992; Ulaby *et al.*, 1986]. En effet, la sensibilité du coefficient de rétrodiffusion à la rugosité augmente, et celle à l'humidité du sol diminue, à mesure que l'angle d'incidence et la fréquence augmentent, et ce, pour toutes les conditions d'humidité du sol [Ulaby et Batlivala, 1976]. Ainsi, une même surface apparaîtra plus rugueuse à des longueurs d'ondes plus courtes et à des angles d'incidence élevés. Ces résultats ont par la suite été obtenus dans de nombreuses études, notamment par Baghdadi *et al.* [2002b] qui ont observé : 1) une augmentation de la sensibilité de σ° à la rugosité de surface avec une augmentation de l'angle d'incidence; 2) qu'à angle d'incidence élevé (47°), l'effet de l'humidité du sol et de l'orientation des rangs des cultures sur σ° est moindre que celui de la rugosité de surface. L'effet de l'angle d'incidence est manifeste lorsque l'on considère qu'en bande C et en polarisation HH, la différence de σ° entre un champ lisse et un autre rugueux est d'environ 9 dB à 47° et de 3 dB à 23° [Baghdadi *et al.*, 2008].

Les études d'Ulaby et Batlivala [1976] et d'Ulaby *et al.* [1978] ont démontré que les effets de la rugosité de surface peuvent être minimisés en utilisant un capteur en polarisation HH ou VV opérant à des fréquences proches de 5 GHz (bande C) à des angles d'incidence entre 7° et 17°. Les données en bande C offrent d'ailleurs le meilleur compromis entre profondeur de pénétration et sensibilité à la rugosité de surface à ces faibles angles d'incidence. Ces propriétés, combinées à la disponibilité des données, expliquent que de nombreuses études portant sur l'estimation de l'humidité du sol utilisent des données radar en bande C.

La rugosité du sol est considérée comme une variation stochastique de la hauteur de la surface du sol par rapport à une surface de référence [Ulaby *et al.*, 1982]. La caractérisation de la rugosité d'une surface est obtenue par l'analyse des variations de hauteur le long de profils mesurant généralement entre 1 et 4 m où l'espacement horizontal entre les mesures de hauteur est entre 1 mm et 2 cm. [Lievens *et al.*, 2009]. Les variations verticales et horizontales sont le plus communément représentées à l'aide de la hauteur quadratique moyenne s (équation (2.6)) et de la longueur de corrélation l [Bryant *et al.*, 2007] :

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2} \quad (2.6)$$

où n est le nombre de mesures de hauteur, z_i la $i^{\text{ème}}$ mesure et \bar{z} la moyenne des mesures

La longueur de corrélation se définit comme étant la distance horizontale pour laquelle l'autocorrélation du profil de surface est supérieure à $1/e$, ou 0,368 [Ulaby *et al.*, 1982; Vulfson *et al.*, 2012]. Les valeurs obtenues sont très variables et dépendent de la longueur des profils utilisés et augmentent asymptotiquement jusqu'à une valeur constante avec une augmentation de la longueur des profils. Comme s et l sont très variables entre les profils réalisés dans un même champ, ils sont habituellement moyennés sur plusieurs profils, généralement entre 3 et 20 [Lievens *et al.*, 2009]. Bryant *et al.* [2007] ont observé qu'au moins 20 profils de 3 m de longueur sont nécessaires à l'obtention d'une mesure de s représentative d'un site d'étude.

Compte tenu de la sensibilité à la longueur d'onde (λ) à laquelle ils sont perçus, s et l sont souvent normalisés à l'aide du nombre d'onde $k = 2\pi/\lambda$ pour donner ks et kl . Ces derniers sont

ensuite utilisés lors de la paramétrisation des modèles d'inversion de l'humidité du sol [Verhoest *et al.*, 2008].

La longueur de corrélation est la mesure la moins précise et la plus difficile à quantifier. En effet, Baghdadi *et al.* [2006a] ont comparé les valeurs simulées par le modèle de l'équation intégrale (IEM) à celles de radar à synthèse d'ouverture (RSO) et ont conclu que les différences étaient liées à l'incertitude des mesures de l et du modèle lui-même. Dans le but de réduire cet écart, des approches empiriques [Baghdadi *et al.*, 2002c; Zribi et Dechambre, 2002], semi-empiriques [Baghdadi *et al.*, 2004] et théoriques [Rahman *et al.*, 2007] permettant de dériver l à partir de mesures de s ont été développées.

Une fonction d'autocorrélation (ACF) qui dépend de la longueur de corrélation est également utilisée pour caractériser la rugosité d'un sol, principalement par le modèle IEM. La sélection de la bonne fonction est importante puisque le coefficient de rétrodiffusion change en fonction de l'ACF choisie [Altese *et al.*, 1996; Fung, 1994]. La fonction exponentielle est la mieux adaptée aux surfaces lisses et est la plus utilisée alors que la fonction Gaussienne est recommandée pour les surfaces rugueuses [Fung, 1994]. La forme de l'ACF est toutefois dépendante de la longueur des profils utilisés pour établir la longueur de corrélation. En effet, pour des sols très rugueux, la forme Gaussienne représente bien la rugosité lorsque les profils sont courts alors que la forme exponentielle la représente mieux lorsqu'ils sont longs.

La caractérisation de la rugosité à l'aide de k_s , k_l et de l'ACF a toutefois ses limites. En effet, plusieurs études ont démontré que le σ° varie de manière importante en fonction de la forme de l'ACF et que la longueur de corrélation est très variable, ce qui introduit des erreurs significatives dans le signal radar modélisé [Altese *et al.*, 1996; Baghdadi *et al.*, 2000, 2004] ainsi que sur les estimations d'humidité du sol.

Les résultats obtenus dans plusieurs études tendent à démontrer que l'utilisation de données acquises dans les mêmes conditions d'humidité du sol à angles d'incidence faible (20°) et élevé (40°) permet d'éliminer les effets de la rugosité et ainsi relier σ° uniquement à l'humidité du sol, en considérant que les conditions de rugosité sont restées les mêmes [Baghdadi *et al.*, 2006b; Srivastava *et al.*, 2003].

Sols végétalisés

La présence de végétation influence de manière importante le coefficient de rétrodiffusion et est probablement la plus importante source d'erreur dans l'estimation de l'humidité du sol. Lorsque de la végétation est présente, la proportion du signal radar pouvant atteindre la surface et ainsi interagir avec l'humidité du sol est influencée de manière importante par celle-ci. La rétrodiffusion totale (σ°) provenant d'un couvert végétal peut-être subdivisée en 3 contributions distinctes : la rétrodiffusion directe du couvert végétal ($\sigma^\circ_{\text{vég}}$), la rétrodiffusion du sol nu ($\sigma^\circ_{\text{surf}}$) et la rétrodiffusion provenant de l'interaction entre le sol et la végétation ($\sigma^\circ_{\text{int}}$) (figure 2.2) et peut être représentée par l'équation (2.7) [Ulaby *et al.*, 1996] :

$$\sigma^\circ = \tau^2 \sigma^\circ_{\text{surf}} + \sigma^\circ_{\text{vég}} + \sigma^\circ_{\text{int}} \quad (2.7)$$

où τ^2 est la transmissivité aller-retour de la végétation, $\sigma^\circ_{\text{surf}}$ la rétrodiffusion du sol nu, $\sigma^\circ_{\text{vég}}$ la rétrodiffusion directe du couvert végétal et $\sigma^\circ_{\text{int}}$ le terme d'interaction entre le sol et le couvert végétal.

L'importance de chacune de ces trois contributions est déterminée par la géométrie et les propriétés diélectriques de la végétation de même que la fréquence, la polarisation et l'angle d'incidence du signal radar [Fung, 1994].

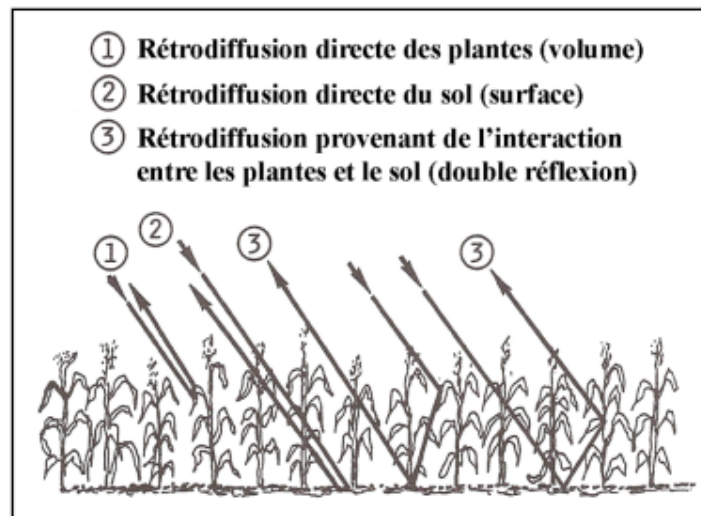


Figure 2.2 Contribution de la rétrodiffusion d'un couvert végétal sur un sol nu. Adaptée de Henderson et Lewis [1998, p. 427]

La sensibilité du signal radar à la végétation en termes de polarisation et de fréquence a été analysée dans de nombreuses études. Les résultats obtenus par Patel *et al.* [2006] démontrent que la sensibilité du signal radar à la densité des plants est la plus élevée en polarisation VH, puis HH et VV, dans l'ordre, autant en bande C que L, et de manière plus marquée en bande L. La plus grande sensibilité de σ°_{VH} à la quantité de végétation est attribuée à la dépolarisation causée par les réflexions multiples se produisant dans le couvert. La plus grande sensibilité en bande L est pour sa part attribuable à sa pénétration du couvert plus importante qu'en bande C, ce qui permet au signal incident d'interagir avec davantage de composantes du couvert. La différence de sensibilité entre les polarisations HH et VV est pour sa part déterminée par l'orientation des branches; des branches horizontales interagiront davantage avec σ°_{HH} et inversement si elles sont plutôt à la verticale [Henderson *et al.*, 1998]. Ainsi, la sensibilité peut être sensiblement la même en polarisation HH et VV, principalement en bande C, compte tenu de la faible pénétration du signal. L'étude d'Ulaby *et al.* [1996] a permis de classifier la végétation haute, courte et les surfaces nues avec une précision de l'ordre de 90% en combinant des données polarimétriques acquises en bande C et L en polarisation HH et HV. Les données radar multifréquence et multipolarisation peuvent donc être utilisées afin d'identifier les zones où les modèles de rétrodiffusion de surface sont applicables.

La sensibilité du signal rétrodiffusé à la végétation en termes d'angle d'incidence est semblable à celle à la rugosité de surface. En effet, plusieurs études ont démontré que la quantité de biomasse perçue, et donc sa contribution au coefficient de rétrodiffusion total, augmentent avec une augmentation de l'angle d'incidence [Baghdadi *et al.*, 2002b; Paloscia, 2002; Romshoo *et al.*, 2002].

Plusieurs études ont été réalisées afin de modéliser la végétation. Les différentes contributions du couvert végétal peuvent être estimées à l'aide de modèles théoriques, notamment ceux développés pour des zones forestières par Chauhan *et al.* [1991] et Karam *et al.* [1992, 1995]. Ces modèles requièrent toutefois l'acquisition de nombreuses données in situ relatives à la géométrie de la végétation (troncs, branches et feuilles), ce qui les rend difficilement utilisables.

Un modèle de diffusion de volume semi-empirique a aussi été développé par Attema et Ulaby [1978]. La végétation y est modélisée comme étant un nuage de gouttelettes d'eau (Water

Cloud Model) et peut être représentée dans le modèle par des variables comme l'indice foliaire (ISF) ou le contenu en eau de la végétation. L'expression dérivée du modèle permet de prédire σ° en fonction de trois paramètres de la cible: l'humidité volumétrique du sol et de la végétation ainsi que la hauteur de la plante. Il ne prend donc pas en compte l'effet des structures solides sur le signal radar. Ainsi, le terme d'interaction $\sigma^\circ_{\text{int}}$ de l'équation (2.7) est jugé négligeable. Le modèle permet aussi d'obtenir une estimation de l'atténuation du signal par la végétation de même que les contributions relatives de la végétation et du sol au coefficient de rétrodiffusion observé. Le faible nombre de paramètres nécessaires à son utilisation, de même que la possibilité d'utiliser des données auxiliaires comme l'ISF, le rendent beaucoup plus facile à utiliser que les modèles théoriques.

Les données polarimétriques ont aussi été utilisées afin de classifier le mécanisme dominant de chaque pixel en trois catégories : réflexion paire, réflexion impaire et réflexion diffuse à l'aide d'un test mathématique [van Zyl, 1989]. Ce principe a par la suite été utilisé par Freeman et Durden [1998] afin d'ajuster un modèle physique, comprenant trois mécanismes de diffusion, aux données polarimétriques elles-mêmes, et ce, sans la nécessité de mesures *in situ*. La décomposition de Freeman–Durden consiste à décomposer le signal radar polarimétrique total en 3 composantes : la puissance de rétrodiffusion de surface (P_s), de volume (P_v) et de double réflexion (P_{DB}), à partir de la matrice de cohérence $[T_{\text{tot}}]$ selon l'équation (2.8)

$$[T_{\text{tot}}] = P_s + P_v + P_{DB} = f_s \begin{bmatrix} 1 & \beta^* & 0 \\ \beta & |\beta|^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{f_v}{4} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + f_d \begin{bmatrix} |\alpha|^2 & \alpha & 0 \\ \alpha^* & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (2.8)$$

$$\text{et} \quad P_s = f_s(1+|\beta|^2) \quad P_v = f_v \quad P_{DB} = f_d(1+|\alpha|^2)$$

où f_s , f_v et f_d sont les amplitudes de diffusion de surface, de volume et de double réflexion, dans l'ordre, et β et α sont des ratios.

Comme la décomposition de Freeman-Durden a été développée pour les conditions présentes en forêt pluvieuse tropicale, des modifications ont été proposées afin d'élargir son domaine de validité [Cui *et al.*, 2012; Hajnsek *et al.*, 2009; Yamaguchi *et al.*, 2005]. L'une des modifications proposées par Hajnsek *et al.* [2009] consiste à remplacer le modèle de diffusion de surface de Bragg par celui de X-Bragg [Hajnsek *et al.*, 2003]. Ce modèle présente trois

avantages, soit : 1) être valide pour des surfaces de forte rugosité, ce qui est le cas en zones agricoles; 2) introduire la diffusion en polarisation croisée, qui permet une meilleure discrimination de la rugosité; et 3) prendre en compte l'effet de la dépolarisation engendrée par la végétation à l'aide du paramètre δ . La matrice de diffusion de surface est alors remplacée par l'équation (2.9).

$$[T_{XB}] = f_s \begin{bmatrix} 1 & \beta^* \sin c(2\delta) & 0 \\ \beta \sin c(2\delta) & \frac{1}{2}|\beta|^2 (1 + \sin c(4\delta)) & 0 \\ 0 & 0 & \frac{1}{2}|\beta|^2 (1 - \sin c(4\delta)) \end{bmatrix} \quad (2.9)$$

et $P_s = f_s (1 + \frac{1}{2}|\beta|^2 (1 + \sin c(4\delta)) + \frac{1}{2}|\beta|^2 (1 - \sin c(4\delta)))$

Les coefficients de rétrodiffusion de surface en polarisation HH ($\sigma_{\text{Surf_HH}}^\circ$) et VV ($\sigma_{\text{surf_VV}}^\circ$) peuvent ensuite être calculés à partir de la décomposition selon les équations (2.10) et (2.11) [Trudel *et al.*, 2012] :

$$\sigma_{\text{Surf_HH}}^\circ = \frac{f_s}{2} + f_s \beta \sin c(2\delta) + \frac{f_s}{4} \beta^2 (1 + \sin c(4\delta)) \quad (2.10)$$

$$\sigma_{\text{Surf_VV}}^\circ = \frac{f_s}{2} - f_s \beta \sin c(2\delta) + \frac{f_s}{4} \beta^2 (1 + \sin c(4\delta)) \quad (2.11)$$

Il a été observé que l'utilisation du modèle de X-Bragg générait des valeurs de diffusion de surface supérieures à celles obtenues avec Bragg, ce qui a permis d'estimer l'humidité du sol sur pratiquement tous les champs agricoles [Trudel *et al.*, 2012]. En effet, la diffusion de surface relative était majoritairement supérieure à 50% pour les pâturages et les champs d'orge, contrairement au modèle de Bragg. De plus, bien que la diffusion de surface ait diminué avec la croissance du maïs, elle est restée près de 50% même lorsque le maïs a atteint 2 m de hauteur [*Ibid.*].

Des analyses de sensibilité ont toutefois permis d'établir que la configuration optimale pour l'estimation de l'humidité du sol sous couvert végétal était l'utilisation d'une fréquence inférieure à 6 GHz et des angles d'incidence inférieurs à 20° [Romshoo *et al.*, 2002; Ulaby *et al.*, 1979]. À mesure que la biomasse augmente, la sensibilité du signal radar aux propriétés du

sol diminue jusqu'à ce que la contribution de l'humidité du sol à la rétrodiffusion totale σ° ne soit plus distinguable.

2.3 Méthodes d'estimation de l'humidité du sol

Il existe de nombreuses méthodes permettant l'estimation de l'humidité du sol. Parmi celles-ci, on compte celles basées sur l'inversion de modèles de rétrodiffusion, la détection de changement, l'optimisation d'une fonction de coût, les interpolations linéaires et non-linéaires, les probabilités statistiques (ex. Bayésiennes), etc. Compte tenu de la popularité des méthodes d'inversion de modèles de rétrodiffusion, cette section traitera des différentes méthodes disponibles. Aussi, une méthode dite d'apprentissage, soit l'approche des réseaux de neurones artificiels, approche employée dans le présent projet, sera abordée de manière plus spécifique. Pour plus de détails sur les autres méthodes, le lecteur est invité à consulter l'article de revue de Kornelson et Coulibaly [2013].

2.3.1 Inversion de modèle de rétrodiffusion

Différents modèles ont été développés pour décrire le coefficient de rétrodiffusion radar (σ°), principalement pour des sols nus ou faiblement végétalisés [Moran *et al.*, 2004] et ultimement inversés afin d'estimer l'humidité du sol. Ces modèles peuvent être classés en trois grandes catégories selon leurs fondements : les modèles empiriques, les modèles semi-empiriques et les modèles théoriques.

Modèles empiriques

Les modèles empiriques tels que ceux développés par Quesney *et al.* [2000], Wang *et al.* [1997] et Zribi et Dechambre [2002] sont basés sur des observations acquises *in situ*. Bien qu'ils puissent estimer l'humidité du sol adéquatement à partir du coefficient de rétrodiffusion, les relations établies pour un site spécifique et une configuration de capteur donnée ne sont généralement pas applicables sur d'autres zones [Dobson *et al.*, 1985; Le Hégarat-Masclé *et al.*, 2002; Moran *et al.*, 2000]. Plusieurs modèles empiriques développés ont la forme générale suivante :

$$\sigma_{dB}^\circ = \alpha m_v + \beta \quad (2.12)$$

où σ_{dB}° est le coefficient de rétrodiffusion en décibel, et α et β sont des constantes estimées empiriquement.

Modèles semi-empiriques

Les modèles semi-empiriques sont basés sur des connaissances théoriques et d'importants ensembles de données de mesures de diffusiomètres réalisées pour des configurations radar et des conditions de surface variées. Ces caractéristiques leur permettent d'être largement applicables. Les plus utilisés sont les modèles de Oh et de Dubois [Dubois *et al.*, 1995a, 1995b; Oh *et al.*, 1992]. Plusieurs mises à jour ont été faites au modèle de Oh afin de prendre en compte les effets de l'autocorrélation de surface [Oh *et al.*, 1994] et l'angle d'incidence [Oh *et al.*, 1994, 2002]. La longueur de corrélation a aussi été retirée des paramètres nécessaires, tandis que la conversion en humidité volumétrique (m_v) a été intégrée en 2004 [Oh, 2004]. Le modèle résultant (équations (2.13) et (2.14)) possède un domaine de validité de $0,04 < m_v < 0,291 \text{ m}^3/\text{m}^3$, $0,13 < ks < 6,98$ pour des angles d'incidence θ entre 10° et 70° .

$$p = \frac{\sigma_{hh}^0}{\sigma_{vv}^0} = 1 - \left(\frac{\theta}{90^\circ} \right)^{0.35 m_v^{-0.65}} e^{-0.4(ks)^{1.4}} \quad (2.13)$$

$$q = \frac{\sigma_{vh}^0}{\sigma_{vv}^0} = 0.095(0.13 + \sin(1.5\theta))^{1.4} (1 - e^{-1.3(ks)^{0.9}}) \quad (2.14)$$

Le modèle de Dubois utilise les données co-polarisées σ_{HH}° et σ_{VV}° en entrée et permet d'inverser simultanément l'humidité du sol et la hauteur quadratique moyenne de la surface selon les équations (2.15) et (2.16) [Dubois *et al.*, 1995a, 1995b]. L'erreur quadratique moyenne (RMSE) obtenue pour m_v est $0,042 \text{ cm}^3/\text{cm}^3$.

$$\sigma_{HH}^\circ = 10^{2,75} \frac{\cos^{1,5} \theta}{\sin^5 \theta} 10^{0,028 \varepsilon \tan \theta} (ks * \sin \theta)^{1,4} \lambda^{0,7} \quad (2.15)$$

$$\sigma_{VV}^\circ = 10^{2,35} \frac{\cos^3 \theta}{\sin^3 \theta} 10^{0,046 \varepsilon \tan \theta} (ks * \sin \theta)^{1,1} \lambda^{0,7} \quad (2.16)$$

où θ est l'angle d'incidence, ε est la partie réelle de la constante diélectrique, s est la hauteur quadratique moyenne de la surface, k est le nombre d'onde et λ la longueur d'onde en cm.

Modèles théoriques

Les modèles théoriques sont basés sur la théorie électromagnétique de Maxwell. Ces modèles simulent la rétrodiffusion à partir des configurations du capteur et des conditions du sol et permettent d'estimer l'humidité du sol par inversion du modèle. Le plus largement utilisé depuis plusieurs décennies est le modèle de l'équation intégrale (IEM) [Fung, 1994]. Selon ce modèle, la contribution de la surface à σ° pour des sols nus en zones agricoles est exprimée par l'équation (2.17):

$$\begin{aligned}\sigma_{pp}^\circ &= \frac{k}{2} |f_{pp}|^2 e^{-4s^2 k^2 \cos^2 \theta} \sum_{n=1}^{\infty} \frac{(4s^2 k^2 \cos^2 \theta)^n}{n!} W^{(n)}(2k \sin \theta, 0) \\ &+ \frac{k^2}{2} |f_{pp}| \operatorname{Re}(f_{pp}^* F_{pp}) e^{-3s^2 k^2 \cos^2 \theta} \sum_{n=1}^{\infty} \frac{(4s^2 k^2 \cos^2 \theta)^n}{n!} W^{(n)}(2k \sin \theta, 0) \\ &+ \frac{k^2}{8} |f_{pp}|^2 e^{-2s^2 k^2 \cos^2 \theta} \sum_{n=1}^{\infty} \frac{(s^2 k^2 \cos^2 \theta)^n}{n!} W^{(n)}(2k \sin \theta, 0)\end{aligned}\quad (2.17)$$

où :

$$f_{hh} = \frac{-2R_h}{\cos \theta}$$

$$f_{vv} = \frac{2R_v}{\cos \theta}$$

$$F_{hh} = 2 \frac{\sin^2 \theta}{\cos \theta} \left[4R_h - \left(1 - \frac{1}{\varepsilon_r}\right) \right] (1 + R_h)^2$$

$$F_{vv} = 2 \frac{\sin^2 \theta}{\cos \theta} \left[\left(1 - \frac{\varepsilon_r \cos^2 \theta}{\mu_r \varepsilon_r - \sin^2 \theta}\right) (1 - R_v)^2 + \left(1 - \frac{1}{\varepsilon_r}\right) (1 + R_v)^2 \right]$$

R_h = coefficient de Fresnel en polarisation horizontale

R_v = coefficient de Fresnel en polarisation verticale

$W^{(n)}$ = la transformée de Fourier de la $n^{\text{ième}}$ puissance du coefficient de corrélation de la surface

ε_r = constante diélectrique

μ_r = permittivité relative

θ = angle d'incidence

s = hauteur quadratique moyenne

k = nombre d'onde

pp = copolarisation (HH ou VV)

Re = partie réelle du nombre complexe

f_{pp}^* = conjugué du nombre complexe f_{pp} .

Les résultats obtenus lors de la modélisation de la rétrodiffusion radar à l'aide du modèle IEM sont très variables. Lors de certaines études, des surestimations sont constamment observées alors que d'autres observent des sous-estimations. Ces erreurs sont souvent associées à une caractérisation inappropriée de la rugosité de surface [Altese *et al.*, 1996; Baghdadi *et al.*, 2006a; Merzouki *et al.*, 2011].

Toutefois, malgré leur domaine de validité étendu, les modèles semi-empiriques et théoriques nécessitent l'acquisition d'une grande quantité de mesures *in situ* relatives à la végétation ainsi qu'à la morphologie du sol, en plus de demander à être calibrés, une opération qui peut être exigeante.

2.3.2 Réseau de neurones artificiels

Dans le but de contourner les limitations des modèles théoriques, empiriques et semi-empiriques, une approche de type non-paramétrique, les réseaux de neurones, a aussi été utilisée pour l'estimation de l'humidité du sol à partir de données RSO [Baghdadi *et al.*, 2002a; Lakhankar *et al.*, 2006; Notarnicola *et al.*, 2008; Pierdicca *et al.*, 2008; Rodriguez-Fernandez *et al.*, 2015; Said *et al.*, 2008; Satalino *et al.*, 2002]. Bien qu'ils ne nécessitent pas de paramètres précis pour être fonctionnels, ils peuvent toutefois être considérés comme des modèles empiriques ou semi-empiriques selon le type de données qu'ils utilisent. En effet, il est possible d'utiliser uniquement des données acquises *in situ* ou de les combiner avec des données obtenues à l'aide de modèles théoriques.

Un réseau de neurones artificiels (RNA) est un système de traitement de données massivement parallélisé possédant certaines caractéristiques des performances ressemblant aux réseaux de neurones biologiques du cerveau humain [Haykin, 1994]. Le développement des RNA est basé sur quatre règles de base 1) le traitement de l'information se produit au niveau de plusieurs éléments appelés neurones, 2) le signal est transmis entre les neurones à l'aide des liens de

connexion, 3) chaque lien est associé à un poids qui représente la force de sa connexion et 4) chaque neurone applique généralement une transformation non-linéaire, appelée fonction d'activation, sur son intrant afin de déterminer son signal sortant [ASCE Task Committee, 2000].

Les RNA sont caractérisés en fonction de leur architecture. Celle-ci peut être définie par le type de connexion, le degré de connectivité et la géométrie adoptée, ce qui détermine le nombre de poids de connexion et la manière selon laquelle l'information s'y propage [Maier et Dandy, 2000]. L'architecture peut être représentée par le tuple (I, N, O, E) où I est un ensemble d'intrant(s), N un ensemble d'unités de calcul (les neurones), O un ensemble de neurone(s) de sortie et E un ensemble de liens directionnels pondérés reliant les neurones entre eux [Rojas, 1996].

2.3.3 RNA et estimation de l'humidité du sol

Au cours des dernières années, de nombreuses études utilisant des RNA afin d'estimer l'humidité du sol ont été réalisées.

Said *et al.* [2008] ont utilisé des données ERS-2 (bande C) en polarisation VV provenant de 3 images acquises à différentes saisons ainsi que différentes combinaisons de données auxiliaires pour entraîner plusieurs RNA. Les jeux de données associés à chacune des images contenaient 112, 102 et 102 données acquises sur sols nus et végétalisés. Les variables entrantes utilisées sont la rugosité de surface s , la couverture du sol, l'ISF, le contenu en eau de la végétation, la hauteur de la végétation, le coefficient de rétrodiffusion calibré (σ°) ou non (VN) et l'angle d'incidence θ° . Au total, 6 combinaisons d'intrants ont été analysées. Ces combinaisons incluaient, en plus de la couverture du sol, σ° et θ° , soit ISF ou s ainsi qu'une ou plusieurs variables liées à la végétation. Les résultats obtenus démontrent que l'utilisation de σ° résulte en des RMSE inférieures et des R^2 supérieurs comparativement à VN. Des coefficients de détermination et des RMSE variant entre 0,84–0,93 et 2,6–6,1%, respectivement, ont été obtenus lors de l'entraînement. L'utilisation de l'ISF comme unique variable liée à la végétation n'a pas entraîné de diminution significative des valeurs de R^2 alors que l'utilisation des données provenant des trois images combinées a produit une diminution marquée des R^2 à 0,75–0,81. Les résultats indiquent aussi que l'exclusion de la rugosité de surface des intrants permet d'obtenir une précision inférieure mais similaire. En effet, des

valeurs de R^2 de 0,87–0,90 avec σ° , couverture du sol, θ° , s et l'ISF contre 0,84–0,89 sans s . Les cartes d'humidité générées à partir des RNA entraînés avec σ° , θ° , couverture du sol et l'ISF comme intrants présentent une variation spatiale de l'humidité du sol en accord avec les conditions observées *in situ* (ex. irrigation, précipitations antécédentes) et une erreur de l'ordre de $\pm 15\%$ comparativement à l'humidité observée tant pour les sols nus que végétalisés.

Les résultats obtenus indiquent donc que des RNA entraînés avec un algorithme de rétro-propagation de l'erreur sont en mesure d'estimer l'humidité du sol avec précision (R^2 de l'ordre de 0,9), et ce, même avec des intrants caractérisant la végétation et la rugosité de surface limités. De plus, des cartes d'humidité du sol avec une précision acceptable ont pu être produites sans utiliser d'information relative à la rugosité de surface. Le principal facteur causant une diminution de la précision des estimations identifié par les auteurs est l'utilisation des données acquises en automne, au printemps et en été. En effet, des RMSE supérieures et des R^2 inférieurs à ceux utilisant les données provenant d'une seule saison ont systématiquement été obtenus lors des phases d'entraînement, de validation et de test, et ce, pour toutes les combinaisons d'intrants. Ce qui, selon les auteurs, pourrait être attribuable à des variations importantes de l'humidités du sol entre les mesures effectuées aux mêmes points d'échantillonnage. Il est aussi avancé que ces variations pourraient être dues au labourage et à la compaction du sol qu'il entraîne, à de l'irrigation ou à des précipitation antérieures.

Une technique utilisée permettant d'entraîner les RNA avec un jeu de données de grande taille représentant des situations variées est l'utilisation de données simulées à partir de modèles théoriques. Notarnicola *et al.* [2008] ont utilisé 1000 des 6000 données radar passives et actives simulées à l'aide des modèles IEM [Fung, 1994] et de Wang [Wang et Choudhury, 1995] ainsi que 79 données acquises *in situ* pour entraîner des RNA pour lesquels la constante diélectrique ϵ était la seule sortie. Deux combinaisons d'intrants de données radar actives ont été retenues et analysées : σ_{HH}° et σ_{VV}° à des angles d'incidence de 23° ou $23^\circ/40^\circ$. Les résultats obtenus démontrent une difficulté à estimer les valeurs de ϵ inférieures à 10 avec seulement les données acquises à un angle d'incidence de 23° comme intrant. En effet, une erreur relative moyenne (MRE) de 36,43 % a été obtenue avec $\sigma_{HH,VV}^\circ$ à 23° contre 23,42 % avec $\sigma_{HH/VV}^\circ$ à $23^\circ/40^\circ$. L'utilisation des données à 23° et $23^\circ/40^\circ$ a permis d'estimer les valeurs de $\epsilon > 10$ avec une précision similaire avec des MRE de 15,68 % et 21,51 %, respectivement. Cette

différence étant due à la faible contribution de l'humidité du sol au signal comparativement à la rugosité et à la végétation pour les sols secs. Les auteurs ont aussi constaté que les RNA surestimaient ε et plus particulièrement les valeurs de $\varepsilon < 10$.

Selon les auteurs, un aspect discutable de la méthode développée est l'utilisation de données synthétiques. En effet, bien qu'elle permette d'obtenir des données représentatives de situations très variées qui ne peuvent être rencontrées expérimentalement, celles-ci ne sont que des représentations approximatives des données réelles. Pour être utiles, elles doivent donc pouvoir être validées à l'aide de nombreuses données expérimentales. De plus, l'écart entre les performances obtenues selon que ε est inférieure ou supérieure à 10 constitue une difficulté importante.

Baghdadi *et al.* [2012b] ont utilisé le modèle IEM afin de générer une base de données contenant 268 110 éléments reproduisant le signal radar pour différentes caractéristiques du sol et du capteur radar. Un RNA a ensuite été entraîné en utilisant les coefficients de rétrodiffusion simulés HH, VV et HV comme paramètres d'entrée et l'humidité du sol et la rugosité de surface (s) comme variables de sortie. Un bruit de $\pm 0.5\text{dB}$ et de $\pm 1\text{ dB}$ a été appliqué sur les simulations de coefficient de rétrodiffusion pour l'entraînement du RNA. L'ajout d'information *a priori* à l'aide de classes d'humidité du sol a permis d'améliorer l'estimation de l'humidité du sol avec le RNA passant d'une RMSE de 6,3 à 4,3 % pour un bruit de $\pm 1\text{dB}$. Dix images RADARSAT-2 en mode polarimétrique à des angles d'incidence de 34° à 47° ont été acquises pour des zones agricoles. L'humidité du sol variait entre 9,0 et 45,7 % lors des acquisitions. L'humidité du sol a ensuite été estimée sur les sols nus ou faiblement végétalisés présent sur les images à l'aide du réseau de neurone. L'ajout d'information *a priori* sur les classes d'humidité du sol a permis de passer d'un r de 0,54 à 0,81 et de 9,8 % à 6,5 % pour la RMSE.

Les résultats obtenus suggèrent donc que l'utilisation d'une importante base de données synthétiques combinée à de l'information *a priori* d'humidité du sol permet d'estimer l'humidité du sol avec une précision acceptable de 6,5 %. Il a aussi été observé que la précision des estimations est indépendante de θ° entre 25° et 45° , ce qui suggère l'absence de corrélation entre les performances des RNA et θ° . La précision obtenue pourrait aussi satisfaire les exigences des produits d'humidité du sol dérivés de la télédétection utilisés de manière

opérationnelle en prévision d'humidité du sol, évaluée à 5 % par Walker *et al.* [2004]. La principale limitation de cette méthode est que l'inversion de l'humidité du sol n'est réalisée que pour les sols nus ou faiblement végétalisés, soit les pixels ayant une valeur de NDVI inférieure à 0,25. Ainsi, l'humidité du sol n'a été obtenue que pour une faible portion du bassin versant à l'étude, ce qui pourrait limiter de manière importante l'utilisation de cette méthode à des fins de modélisation ou de prévision hydrologique.

Plus récemment Baghdadi *et al.* [2016] ont entraîné des RNA possédant une couche cachée contenant 20 neurones à l'aide de données provenant de 18 images RADARSAT-2 acquises en mode polarimétrique et autant d'images LANDSAT-7/8 afin d'estimer m_v et l'index de surface foliaire (ISF). Les bénéfices de l'utilisation de données RSO en double polarisation ou complètement polarimétriques, comparativement aux données en simple polarisation, ont été évalués. L'utilisation des paramètres polarimétriques de l'entropie de Shannon (SE) et des composantes de Pauli a aussi été étudiée. Au total, six combinaisons d'intrants incluant toutes l'angle d'incidence ont été testées. Les résultats obtenus démontrent que le RNA utilisant σ°_{HH} et le NDVI a permis les meilleures estimations de m_v . De plus, l'utilisation d'autres indices de végétation, de données en polarisation double ou quadruple ou l'ajout de paramètres polarimétriques a eu peu d'influence. En effet, une erreur relative moyenne (MRE) de 20,3 % a été obtenue avec σ°_{HH} et le NDVI comme intrants contre près de 26 % avec σ°_{HV} ou σ°_{VH} et NDVI et 21 % environ avec l'ajout des paramètres polarimétriques. Les RNA utilisant θ° , σ°_{HH} et le NDVI ont été en mesure d'estimer m_v sur l'ensemble du domaine de valeur (10–45 %) alors que l'ajout de σ°_{HV} ou σ°_{VV} a produit une surestimation des valeurs d'humidité du sol entre 10 et 25 % et une sous-estimation pour les valeurs entre 25 et 40 %. Les analyses de l'erreur en fonction des valeurs de NDVI, de la biomasse, du contenu en eau et de la hauteur de la végétation indiquent que la contribution du sol à σ° reste importante même lorsque le couvert est dense mais que l'erreur augmente avec la densité. Les résultats ont aussi confirmé que la contribution du sol à σ° est faible à angles d'incidence élevés. En effet, l'erreur quadratique moyenne obtenue sur m_v était de 5,2 % pour les angles d'incidence entre 22° et 43° contre 9,2 % à 47°.

Les résultats obtenus dans cette étude suggèrent donc que l'ajout d'une ou plusieurs variables moins sensibles à l'humidité du sol que σ°_{HH} , à l'exception du NDVI, n'améliore pas significativement la capacité des RNA à estimer l'humidité du sol. Les variables étudiées

provenant d'observations satellites incluent σ_{VV}° , σ_{HV}° , les paramètres polarimétriques de l'entropie de Shannon et les composantes de Pauli alors que la fraction du rayonnement photosynthétiquement actif absorbée et la fraction de couverture végétale provenaient de mesures *in situ*. L'influence du NDVI comme variable entrante aux RNA sur les estimations de l'humidité du sol n'est toutefois pas un aspect abordé au cours de cette recherche puisque l'ISF était l'une des sorties. Il est donc possible que l'utilisation de σ_{HH}° comme seul intrant autre que θ° ait donné de meilleurs résultats. L'objectif de cette étude étant d'évaluer les capacités d'estimer l'humidité du sol uniquement sur des sols végétalisés, plus précisément des prairies irriguées, les performances des RNA développés sur des sols nus ne sont pas connues. Par conséquent, la méthode développée est applicable spécifiquement pour ce type de couverture du sol.

CHAPITRE 3 SITE D'ÉTUDE ET DONNÉES

Ce chapitre présente la zone d'étude de même que l'ensemble des données utilisées au cours de cette recherche. Les données caractérisant les bassins versants à l'étude, les données acquises *in situ* ainsi que les données de télédétection y sont décrites.

3.1 Zone d'étude

Le territoire à l'étude est montré à la figure 3.1. Il est situé dans la sous-région Centre de l'Ontario et est constitué de deux bassins versants adjacents, soit ceux des rivières Trent, à l'est, et Severn à l'ouest. Bien qu'ils fassent tous deux partie du bassin hydrographique des Grands Lacs et du Saint-Laurent, ils constituent deux bassins versants distincts. En effet, la rivière Trent est un tributaire du lac Ontario alors que la rivière Severn se jette dans le lac Huron.



Figure 3.1 Localisation de la zone d'étude

Les bassins versants Trent et Severn drainent des superficies de 12 200 km² et 6 160 km², respectivement, pour un total de 18 360 km² [Ecoplans Limited, 2007]. Les limites

géographiques de la zone d'étude sont les latitudes 44°00'00" N et 45°22'00" N et les longitudes 77°21'00" O et 79°47'00" O.

La zone d'étude est traversée par la voie navigable Trent-Severn reliant Trenton et Port Severn sur une longueur de 386 km. Elle est composée de canaux intérieurs, de rivières et de lacs naturels reliés entre eux par des passes, des sas et des barrages. Dans le secteur des hautes terres d'Haliburton, dans le nord du bassin versant de la rivière Trent, des barrages ont été construits sur 44 lacs. L'eau de ruissellement printanière y est stockée puis relâchée au cours de l'été afin d'approvisionner la portion de la voie navigable située dans le bassin versant Trent. Au total, la voie navigable comprend 160 barrages et ouvrages de contrôle des débits en plus de 45 écluses et 1700 aides à la navigation [Parks Canada, 2011]. Les bassins versants Trent et Severn sont donc des bassins où les débits sont fortement régulés.

Les bassins versants Trent et Severn s'étendent sur deux régions physiographiques : le Bouclier canadien, plus précisément la région laurentienne, qui couvre environ les deux tiers septentrionaux, et les basses-terres du Saint-Laurent au sud. Les dépôts meubles sont en majorité inexistantes dans la partie nord du Bouclier et inférieures à 10 m dans sa partie sud-est. Le dépôt s'épaissit graduellement vers le sud pour atteindre 262 m à l'extrême sud et à l'ouest du lac Simcoe (annexe A.1).

Les matériaux dominants du dépôt de surface au sud du bassin versant sont le till, le sable, le limon et le gravier, en ordre d'importance. La zone de transition entre les basses-terres du Saint-Laurent et le Bouclier canadien est dominée par le till alors qu'au nord, le roc domine le gravier et le sable présents dans une faible proportion (annexe A.2). Une carte représentant la perméabilité en fonction de ces matériaux est présentée à l'annexe A.3. Une grande partie de la portion nord des bassins versants Trent et Severn repose majoritairement sur la roche mère et l'épaisseur du dépôt de surface est inférieure à 25 cm dans l'ensemble du bassin versant à l'exception d'une portion des terres entourant le lac Simcoe, où elle atteint 1 m (annexe A.4).

Les sols s'étant développés sur la zone située dans le Bouclier canadien sont majoritairement non-classifiés (dominés par le roc), mais les brunisols mélaniques et dystriques ainsi qu'une petite zone de podzol humo-ferrique y sont aussi présents. Les basses-terres du Saint-Laurent sont quant à elles principalement composées de brunisol mélanique au nord et de luvisol brun gris au sud. On retrouve aussi trois zones distinctes : des gleysols au nord-est du lac Simcoe et

des sols d'origine organique humisol au sud de ce dernier. La carte des ordres et grands groupes de sols est présentée à l'annexe A.5. Les sols peu épais du Bouclier canadien ne permettent que le stockage d'une quantité limitée d'eau, ce qui entraîne un écoulement rapide. Les sols épais des basses-terres du Saint-Laurent peuvent quant à eux accumuler des quantités importantes d'eau, ce qui leur permet de retenir et relâcher plus lentement les eaux provenant des précipitations et de la fonte printanière. La mise à nu du sol à des fins agricoles diminue toutefois de manière importante cette capacité de rétention d'eau et augmente le ruissellement de surface.

La partie nord est principalement constituée de forêts feuillues et mixtes contrairement à la partie sud, plus densément peuplée et constituée d'une mosaïque de terres cultivées, de zones forestières, de milieux humides et de zones développées (annexe A.6).

Le climat ontarien peut, de manière générale, être considéré comme continental humide et est affecté par trois sources d'air : l'air sec et polaire provenant du nord, facteur dominant durant l'hiver, l'air polaire du Pacifique arrivant des prairies de l'ouest et l'air subtropical chaud et humide de l'océan Atlantique et du Golfe du Mexique [Baldwin *et al.*, 2011]. L'influence de ces flux d'air sur la température et les précipitations est dépendante de la latitude, de la proximité de masses d'eau importantes et de la topographie, dans une moindre mesure [*Ibid.*]. La température annuelle moyenne varie entre 3,5 et 7,2°C selon un gradient nord-sud. La température minimale moyenne hivernale se situe entre -9,3°C et -16°C et la température maximale moyenne estivale entre 22,3°C et 24,7°C. Le même gradient thermique est observé tout au long de l'année mais la proximité du lac Ontario apporte des hivers plus doux et des étés plus frais dans la portion centrale du sud du bassin versant comparativement aux zones sud-ouest et sud-est. L'amplitude thermique est beaucoup plus importante en hiver qu'en été (7°C comparativement à 2,5°C). Les précipitations annuelles moyennes varient de 813 mm à 1105 mm selon un gradient nord-ouest–sud-est. En hiver, la portion nord-ouest du bassin versant reçoit plus de précipitations, et ce, de manière marquée, alors qu'il en tombe nettement moins au sud du lac Simcoe. En été, le nord est la zone où il pleut le plus abondamment alors que le sud-est est celle recevant le moins de précipitations. Les cartes de températures et de précipitations sont présentées à l'annexe A.7.

3.2 Données sur les bassins versants Trent et Severn

Afin d'être en mesure de prendre en considération certaines variables influençant le coefficient de rétrodiffusion radar ainsi que les conditions d'humidité du sol, différentes données auxiliaires ont été utilisées. Ces données sont les données liées à la granulométrie du sol minéral, à la pente du terrain et à la quantité de biomasse.

3.2.1 Données de texture du sol

Les fractions de sable, de limon et d'argile du sol ont été extraites des données de pédo-paysage du Canada (PPC) [Le groupe de travail sur les Pédo-paysages du Canada, 2010]. Ces données sont produites par Agriculture et Agroalimentaire Canada et diffusées par le Service d'Information sur les Sols du Canada (SISCan). Elles contiennent l'ensemble des données sur les sols arables du Canada à une échelle de 1 : 1 000 000. Les PPC se fondent sur les cartes pédologiques existantes recompilées selon une échelle de 1 / 1 million [ibid]. Les cartes pédologiques de l'Ontario utilisées lors de la recompilation pour les comtés de Simcoe, York, Kawartha Lakes et Peterborough sont à une échelle de 1 : 63 360 et celui de Durham 1 : 126 720. L'information pour les comtés de Muskoka et Halliburton n'est pas disponible. La carte de texture du sol originale est présentée à l'annexe A.8. TEXTURE DU SOL

3.2.2 Données de pente

La pente du terrain a été dérivée de modèles numériques d'élévation (MNÉ). Les MNÉ utilisés sont les données numériques d'élévation du Canada (DNEC) à l'échelle 1:50 000 distribuées par GéoBase puis GéoGratis par la suite [Gouvernement du Canada, 2007]. Au total, 30 jeux de données ont été utilisés. Ces données ont été produites par le Ministère des richesses naturelles de l'Ontario à l'aide de sources aux échelles 1/10 000 et 1/20 000. La résolution spatiale des pixels est de 0,75 seconde d'arc, ce qui correspond à 23m x 16m en latitude et en longitude respectivement, et la résolution verticale est de 1 m. La précision d'alignement horizontal est de 5 mètres alors qu'elle est de 3 m sur le plan vertical, à l'exception d'un jeu de données. Dans la zone d'étude, excluant les lacs, les pentes sont inférieures à 10 % pour 84,6% des pixels. L'histogramme de fréquence de pente et les caractéristiques des jeux de données des MNÉ sont présentées aux annexes b et ANNEXE C, respectivement.

3.3 Données *in situ*

3.3.1 Données d'humidité du sol

Au total, 40 sondes semi-permanentes ont été installées et fixées à un tuteur. Les emplacements, présentés à la figure 3.2, ont d'abord été conditionnés par une analyse multicritère basée sur la topographie, le type de sol, l'occupation du sol, et dans une perspective plus pratique, par l'accessibilité et la disponibilité des sites potentiels pouvant héberger les sondes durant toute la période d'acquisition des données. Elles sont distribuées essentiellement dans le 2/3 sud des bassins versants. Les sols du tiers nord étant essentiellement sous couvert forestier, uniquement trois sondes y ont été installées.



Figure 3.2 Carte de localisation des sondes d'humidité du sol et des stations météorologiques

Les mesures ont été effectuées à l'aide de sondes d'humidité du sol EC-5 de Decagon (figure 3.3). Ces sondes mesurent la constante diélectrique du sol à partir de la capacité électrique en

utilisant la technologie *Frequency Domain* pour des températures entre -40°C et 60°C . Les mesures obtenues ont une précision de $\pm 0,03 \text{ m}^3/\text{m}^3$ et une résolution de $0,001 \text{ m}^3/\text{m}^3$ [Decagon Devices, Inc., 2001]. Notons que l'effet de la texture et de la salinité du sol est minimisé par l'utilisation d'une fréquence de 70 MHz.

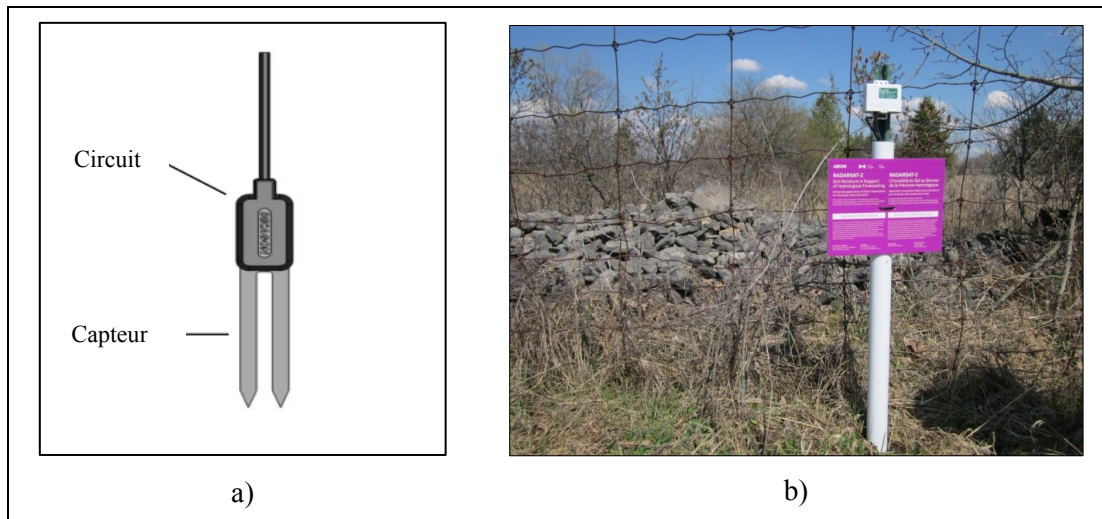


Figure 3.3 a) Diagramme de la sonde EC-5 (adapté de Decagon Devices, Inc. [2001]) et b) station de mesure

L'humidité volumétrique du sol (m_v) en m^3/m^3 (note : pour l'eau, $1\text{g}/\text{cm}^3 = 1 \text{ m}^3/\text{m}^3$) a été mesurée à une profondeur de 5 cm aux 2 heures sur 38 et 34 sites entre avril et octobre 2012 et 2013, respectivement. En 2012, une sonde a été endommagée par un rongeur et n'a pas été remplacée alors qu'en 2013, une des sondes a disparu et une seconde a été retrouvée brisée. Enfin, les données d'une et de trois sondes, en 2012 et 2013 respectivement, étaient invalides et n'ont donc pas été utilisées, ce qui explique la différence entre le nombre de sites pour lesquels des enregistrements sont disponibles et le nombre de sondes installées. Au cours des saisons 2012 et 2013, 119, 208 et 534 couples sondes/images ont été obtenus pour les images RADARSAT-2 acquises en mode polarimétrique, standard/large et combinés, dans l'ordre. Une description de ces modes (résolution, couverture, angle d'incidence) est présentée à la section 3.4.2. Les valeurs de m_v mesurées au moment de l'acquisition des images RADARSAT-2 variaient entre 0,07 et 0,48 m^3/m^3 , avec une moyenne de 0,25 m^3/m^3 et un écart-type de 0,09 m^3/m^3 . La dispersion des valeurs de m_v en fonction du mode d'acquisition des images est présentée à la figure 3.4.

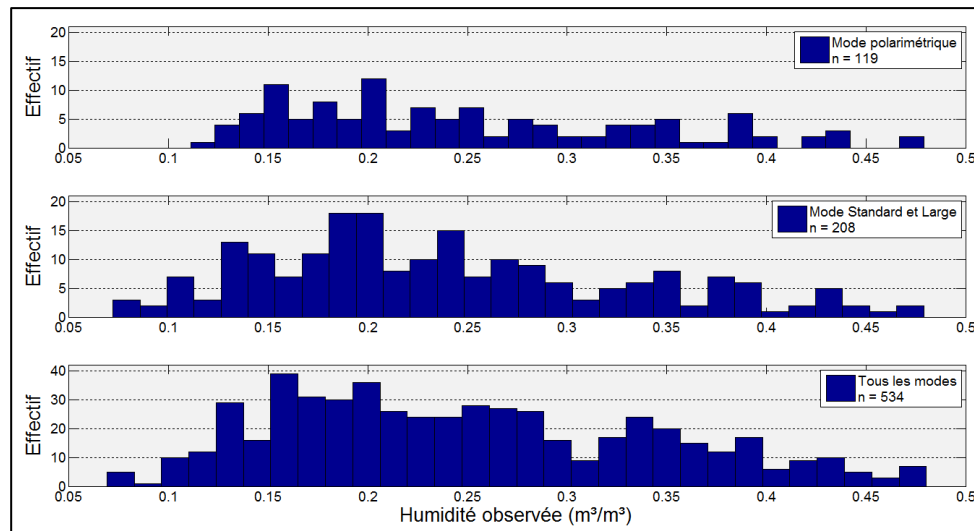


Figure 3.4 Dispersions des données d'humidité du sol observée

3.3.2 Données météorologiques

Les données météorologiques journalières de 2012 et 2013 ont été téléchargées pour 11 stations météorologiques d'Environnement et Changement climatique Canada [Gouvernement du Canada, 2016]. Les fichiers obtenus contiennent de nombreuses données mais seules les précipitations totales ont été utilisées. Les données de précipitations ont été utilisées pour dresser un portrait des conditions présentes lors de l'acquisition des images RADARSAT-2 et celles des jours précédents.

3.4 Données de télédétection

3.4.1 Données optiques

La végétation a été caractérisée, entre autres, à l'aide de l'indice de surface foliaire (ISF). L'ISF est un paramètre adimensionnel représentant l'aire totale des faces supérieures des feuilles par unité de surface au sol.

Au total, 23 images d'ISF composites sur huit jours obtenues du capteur *Moderate Resolution Imaging Spectroradiometer* (MODIS) d'une résolution spatiale de 1 km et sur un domaine de valeurs de [0,10] ont été utilisées. Les images couvrant la zone d'étude correspondent à la tuile h12v4 (horizontale 12, verticale 4) du système de coordonnées MODIS et ont été téléchargées du site internet *Land Processes Distributed Active Archive Center* (LP DAAC) de la NASA. Le tableau 3.1 présente les images d'ISF acquises ainsi que la date de début et de fin de la période de composition.

Tableau 3.1 Images d'ISF acquises

Nom du fichier	Date de début	Date de fin
MOD15A2.A2012097.h12v04.005.2012108170313.hdf	2012-04-06	2012-04-13
MOD15A2.A2012105.h12v04.005.2012114152130.hdf	2012-04-14	2012-04-21
MOD15A2.A2012113.h12v04.005.2012122080251.hdf	2012-04-22	2012-04-29
MOD15A2.A2012121.h12v04.005.2012130083815.hdf	2012-04-30	2012-05-07
MOD15A2.A2012129.h12v04.005.2012141234842.hdf	2012-05-08	2012-05-15
MOD15A2.A2012137.h12v04.005.2012146150034.hdf	2012-05-16	2012-05-23
MOD15A2.A2012145.h12v04.005.2012156153552.hdf	2012-05-24	2012-05-31
MOD15A2.A2012153.h12v04.005.2012166170255.hdf	2012-06-01	2012-06-08
MOD15A2.A2012161.h12v04.005.2012170080220.hdf	2012-06-09	2012-06-16
MOD15A2.A2012169.h12v04.005.2012178103921.hdf	2012-06-17	2012-06-24
MOD15A2.A2012185.h12v04.005.2012209002301.hdf	2012-07-03	2012-07-10
MOD15A2.A2012193.h12v04.005.2012202172441.hdf	2012-07-11	2012-07-18
MOD15A2.A2012217.h12v04.005.2012229064804.hdf	2012-07-19	2012-07-26
MOD15A2.A2013129.h12v04.005.2013141204503.hdf	2013-05-09	2013-05-16
MOD15A2.A2013137.h12v04.005.2013146125525.hdf	2013-05-17	2013-05-24
MOD15A2.A2013145.h12v04.005.2013155101751.hdf	2013-05-25	2013-06-01
MOD15A2.A2013153.h12v04.005.2013162141217.hdf	2013-06-02	2013-06-09
MOD15A2.A2013161.h12v04.005.2013170142153.hdf	2013-06-10	2013-06-17
MOD15A2.A2013169.h12v04.005.2013178193209.hdf	2013-06-18	2013-06-25
MOD15A2.A2013177.h12v04.005.2013186131311.hdf	2013-06-26	2013-07-03
MOD15A2.A2013193.h12v04.005.2013204214546.hdf	2013-07-12	2013-07-19
MOD15A2.A2013201.h12v04.005.2013213181826.hdf	2013-07-20	2013-07-27
MOD15A2.A2013209.h12v04.005.2013218194807.hdf	2013-07-28	2013-08-04

L'utilisation de données d'ISF ayant une résolution de 1 km pour la production de cartes d'humidité du sol à une résolution de 12,5 m n'est toutefois pas idéale. En effet, les variations de la quantité de biomasse, et donc de l'ISF, à l'échelle du champ ne peuvent être prises en compte par les données MODIS. Cette lacune risque donc d'apporter de l'incertitude lors de la modélisation de la végétation, et ce, principalement dans les zones de végétation hétérogène. Le fait qu'il s'agisse d'images composites 8 jours augmente aussi la période nécessaire à l'observation d'une variation notable de l'ISF. En effet, l'écart entre les valeurs de 2 images consécutives risque fort d'être minime. Ces compromis ont été jugés acceptables puisque la composition sur 8 jours permet d'obtenir, en une seule image, des images libres de nuages et de leurs ombres sur l'ensemble de la zone d'étude.

3.4.2 Données radar à synthèse d'ouverture (RSO)

Un total de 47 images radar en bande C (fréquence de 5,405 GHz) captées par RADARSAT-2 entre avril et août 2012 et entre mai et juillet 2013 ont été acquises. Ces images ont été fournies par l'Agence spatiale canadienne (ASC) dans le cadre du Programme de Développement d'Applications en Observation de la Terre (PDAOT) pour le projet « RADARSAT-2, l'humidité du sol au service de la prévision hydrologique ». De ces images, 30 ont été acquises en mode Quad-pol standard (SQ) en polarisation HH, VV, HV et VH, 7 en mode standard (S) en polarisation HH, 10 en mode Large (W), dont 7 en polarisation HH et 3 en HH et HV. Les spécifications de ces images selon les informations fournies par MacDonald, Dettwiler and Associates Ltd. [2014] sont présentées au tableau 3.2. Les images acquises en mode standard, large, et Quad-pol standard ont une couverture nominale de 100 x 100 km, 150 x 150 km et 25 x 25 km, respectivement. Les empreintes au sol des images acquises sont présentées à la figure 3.5.

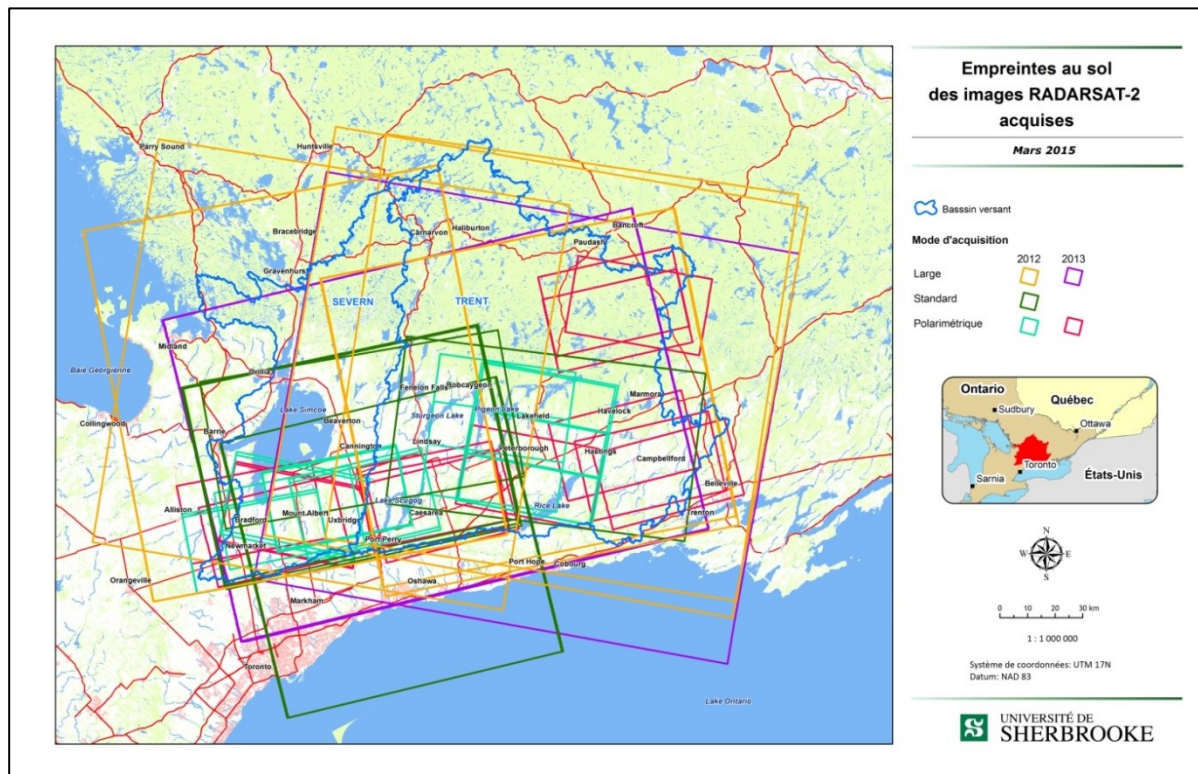


Figure 3.5 Empreintes au sol des images RADARSAT-2 acquises

Tableau 3.2 Sommaire des caractéristiques des images RADARSAT-2 acquises

Date	Polarisations	Faisceau	Résolution nominale (m) ¹	Heure (UTC)	Orbite	Angle d'incidence (°)
4/10/2012	HH VV HV VH	SQ18W	13,5 x 7,6	23:11:52	Ascendant	37 - 39
4/11/2012	HH	W2	13,5 x 7,7	11:22:49	Descendant	31 - 40
4/14/2012	HH VV HV VH	SQ1W	9,0 x 7,6	11:35:29	Descendant	18 - 21
4/14/2012	HH VV HV VH	SQ1W	9,0 x 7,6	11:35:31	Descendant	18 - 21
4/17/2012	HH	S3	13,5 x 7,7	23:07:41	Ascendant	31 - 37
4/18/2012	HH VV HV VH	SQ21W	13,5 x 7,6	11:18:50	Descendant	40 - 42
4/18/2012	HH VV HV VH	SQ21W	13,5 x 7,6	11:18:53	Descendant	40 - 42
4/21/2012	HH	W1	13,5 x 7,7	11:31:08	Descendant	19 - 21
4/25/2012	HH	S6	13,5 x 7,7	11:14:41	Descendant	41 - 47
4/27/2012	HH VV HV VH	SQ21W	13,5 x 7,6	23:16:01	Ascendant	40 - 42
4/28/2012	HH	W2	13,5 x 7,7	11:27:02	Descendant	31 - 40
5/1/2012	HH VV HV VH	SQ1W	9,0 x 7,6	11:39:47	Descendant	18 - 21
5/1/2012	HH	S1	9,0 x 7,7	22:59:17	Ascendant	20 - 27
5/4/2012	HH	W3	13,5 x 7,7	23:11:51	Ascendant	39 - 45
5/8/2012	HH VV HV VH	SQ1W	9,0 x 7,6	11:35:31	Descendant	18 - 21
5/8/2012	HH VV HV VH	SQ1W	9,0 x 7,6	11:35:34	Descendant	18 - 21
5/11/2012	HH	S3	13,5 x 7,7	23:07:44	Ascendant	31 - 37
5/21/2012	HH	W3	13,5 x 7,7	23:16:02	Ascendant	39 - 45
5/25/2012	HH VV HV VH	SQ1W	9,0 x 7,6	11:39:47	Descendant	18 - 21
5/25/2012	HH	S1	9,0 x 7,7	22:59:16	Ascendant	20 - 27
5/28/2012	HH	W3	13,5 x 7,7	23:11:49	Ascendant	39 - 45
6/4/2012	HH	S3	13,5 x 7,7	23:07:42	Ascendant	31 - 37
6/7/2012	HH	S7	13,5 x 7,7	23:20:14	Ascendant	44 - 49
6/14/2012	HH	W3	13,5 x 7,7	23:16:01	Ascendant	39 - 45
6/21/2012	HH VV HV VH	SQ18W	13,5 x 7,6	23:11:52	Ascendant	37 - 39
7/8/2012	HH VV HV VH	SQ21W	13,5 x 7,6	23:16:00	Ascendant	40 - 42
7/12/2012	HH VV HV VH	SQ4W	9,0 x 7,6	22:59:23	Ascendant	22 - 25
7/15/2012	HH VV HV VH	SQ18W	13,5 x 7,6	23:11:52	Ascendant	37 - 39
8/5/2012	HH VV HV VH	SQ4W	9,0 x 7,6	22:59:23	Ascendant	22 - 25
8/8/2012	HH VV HV VH	SQ17W	13,5 x 7,6	23:11:52	Ascendant	36 - 38
5/16/2013	HH VV HV VH	SQ21W	13,5 x 7,6	23:16:02	Ascendant	40 - 42
5/20/2013	HH VV HV VH	SQ1W	9,0 x 7,6	11:39:46	Descendant	18 - 21
5/23/2013	HH VV HV VH	SQ20W	13,5 x 7,6	23:11:51	Ascendant	39 - 41
5/27/2013	HH VV HV VH	SQ5W	9,0 x 7,6	22:55:13	Ascendant	23 - 26
6/3/2013	HH HV	W1	13,5 x 7,7	11:31:11	Descendant	19 - 31
6/3/2013	HH HV	W1	13,5 x 7,7	11:31:11	Descendant	19 - 31
6/6/2013	HH VV HV VH	SQ15W	13,5 x 7,6	23:03:40	Ascendant	34 - 37
6/13/2013	HH VV HV VH	SQ2W	9,0 x 7,6	22:59:23	Ascendant	19 - 23
6/16/2013	HH VV HV VH	SQ20W	13,5 x 7,6	23:11:51	Ascendant	39 - 41
6/20/2013	HH VV HV VH	SQ1W	9,0 x 7,6	22:55:13	Ascendant	18 - 21
6/27/2013	HH VV HV VH	SQ1W	9,0 x 7,6	22:51:02	Ascendant	18 - 21
6/30/2013	HH VV HV VH	SQ15W	13,5 x 7,6	23:03:31	Ascendant	34 - 37
6/30/2013	HH VV HV VH	SQ15W	13,5 x 7,6	23:03:39	Ascendant	34 - 37
7/14/2013	HH VV HV VH	SQ1W	9,0 x 7,6	11:35:32	Descendant	18 - 21
7/21/2013	HH VV HV VH	SQ4W	9,0 x 7,6	11:31:13	Descendant	22 - 25
7/31/2013	HH VV HV VH	SQ1W	9,0 x 7,6	11:39:46	Descendant	18 - 21
7/31/2013	HH HV	W2	13,5 x 7,7	22:59:20	Ascendant	19 - 31

¹Distance-temps x azimuth

Bien que les faibles angles d'incidence soient mieux adaptés pour l'estimation de l'humidité du sol, 25 images ont été acquises à angles d'incidence élevés. Cette situation résulte de plusieurs conflits d'usage liés à des acquisitions situées à proximité de la zone d'étude ayant un niveau de priorité supérieur à celui de la présente étude. Parmi les images acquises, 18 l'ont été en orbite descendante, ce qui peut avoir un impact important dans le cadre d'une étude portant sur l'humidité du sol. En effet, la rosée matinale peut affecter le signal radar de manière importante. Toutefois la rosée ne pose pas problème dans notre cas puisque l'acquisition des images a été faite entre 10:45 et 11:45, où le risque de présence de rosée est faible.

CHAPITRE 4 MÉTHODOLOGIE

Ce chapitre présente la méthodologie développée au cours du projet. La réalisation de ce projet a été effectuée en 8 grandes étapes telles qu'illustrées dans l'organigramme de la figure 4.1. Les étapes liées au traitement des images RADARSAT-2 sont présentées dans l'organigramme de la figure 4.2 alors que celles relatives aux RNA sont présentées à l'organigramme de la figure 4.3.

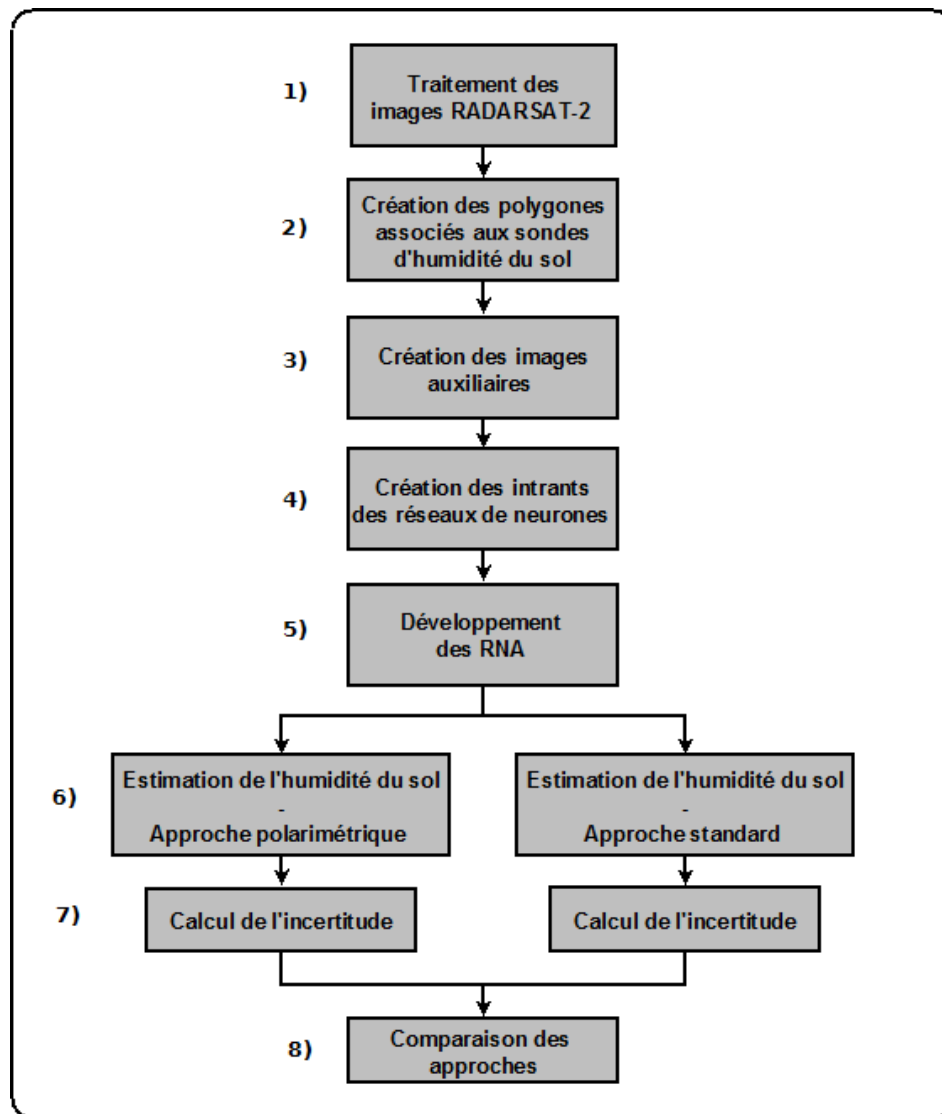
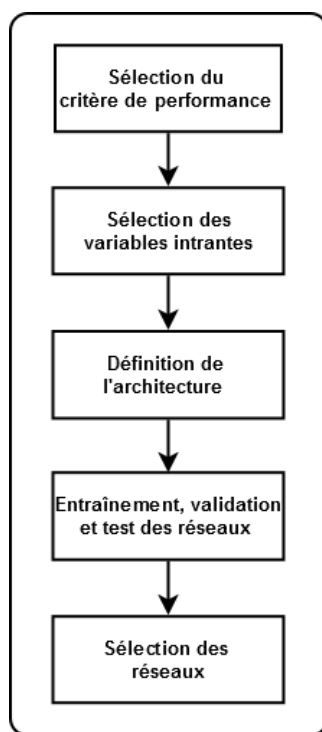
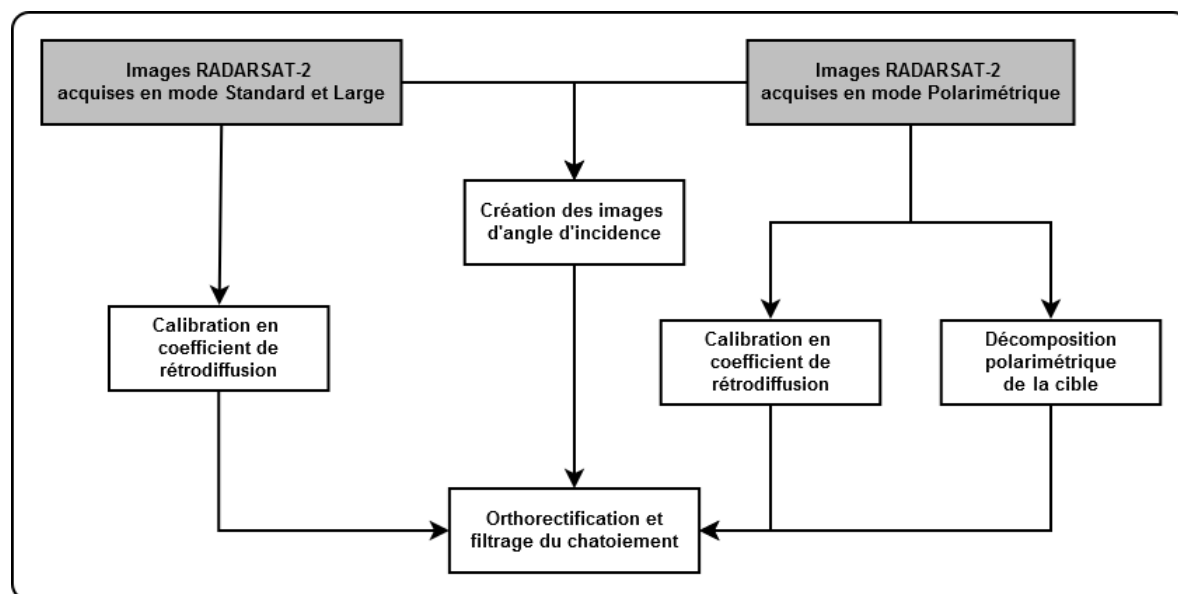


Figure 4.1 Organigramme méthodologique global



4.1 Traitement des images RSO

Dans un premier temps, les images livrées ont été ouvertes dans le logiciel de traitement d'images Geomatica de la firme PCI [PCI Geomatics, 2009] afin de créer des images multi bandes regroupant les données de phase et d'amplitude dans un fichier en format .pix.

4.1.1 Création des images d'angle d'incidence

Afin d'être en mesure d'obtenir une image contenant les angles d'incidence, un fichier texte contenant la valeur de l'angle d'incidence pour chacune des colonnes de l'image a été créé pour chacune des images RADARSAT-2 acquises. Ces valeurs ont été calculées à partir des valeurs des angles d'incidence à la portée proximale et distale (*near range* et *far range*) et du nombre d'échantillons en portée contenus dans le fichier de métadonnées associé à chacune des images RADARSAT-2. Cette étape a été réalisée à l'aide de PCI Geomatica [PCI Geomatics, 2009]. Une fonction Matlab® utilisant le nombre de lignes et de colonnes présentes dans chaque image ainsi que le fichier d'angles d'incidence a été développée pour générer l'image finale. L'image obtenue a ensuite été ajoutée comme une nouvelle bande aux images brutes dans PCI Geomatica.

4.1.2 Étalonnage radiométrique des images RSO

Cette étape a été réalisée pour l'ensemble des images acquises. Elle consiste à extraire le coefficient de brillance radar (β°) à partir des valeurs numériques (VN) des données brutes, puis à le convertir en coefficient de rétrodiffusion radar (σ°). Les données utilisées sont en format *Single Look Complex*. Elles sont donc composées d'une partie réelle (I) et d'une partie imaginaire (Q) représentant respectivement l'amplitude et la phase. La valeur de β° pour la $j^{\text{ième}}$ rangée de pixels est donnée par l'équation (4.1) :

$$\beta_j^\circ = 10 * \log_{10} \left[\left(\frac{VNI_j}{A2_j} \right)^2 + \left(\frac{VNQ_j}{A2_j} \right)^2 \right] dB \quad (4.1)$$

où VNI_j et VNQ_j sont les valeurs numériques des composantes I et Q du $j^{\text{ième}}$ pixel à partir du début de la ligne de portée et $A2_j$ est la portée correspondante en fonction du gain.

La relation permettant de convertir β° en σ° est donnée par l'équation (4.2) [RADARSAT International, 2000] :

$$\sigma_j^o = \beta_j^o + 10 * \log_{10}(\sin I_j) dB \quad (4.2)$$

où I_j est l'angle d'incidence au pixel de portée j . Cette formule repose sur l'hypothèse que la Terre est un ellipsoïde régulier au niveau de la mer.

L'étalonnage décrit dans la présente section (équations 4.1 et 4.2) a été réalisé à l'aide du logiciel de traitement d'image libre NEST 4C, distribué par l'Agence spatiale européenne [ESA, 2012]. Les images de coefficient de rétrodiffusion obtenues ont été ajoutées comme de nouvelles bandes aux images brutes à l'aide de PCI Geomatica.

4.1.3 Décomposition polarimétrique

Une décomposition polarimétrique de la cible de Freeman-Durden [Freeman et Durden, 1998] est appliquée sur les 30 images brutes acquises en mode Quad-pol standard. Cette variante de décomposition du signal radar a été préférée aux autres méthodes pour les raisons suivantes: 1) elle est valide pour des surfaces de forte rugosité comme les champs agricoles, 2) les résultats obtenus ont une explication physique claire, ce qui en facilite l'analyse et l'interprétation visuelle, 3) elle permet d'isoler la contribution de la surface à σ^o , 4) elle ne nécessite pas de mesures *in situ* et 5) elle a été utilisée par plusieurs auteurs dans des études portant sur l'estimation de l'humidité du sol (*ex. Hajnsek et al., 2009; Trudel et al., 2012*).

Cette procédure a été réalisée en deux étapes. La première étape est réalisée à l'aide du logiciel libre PolSARpro [ESA, 2011] et consiste à extraire les matrices de diffusion 2x2 [S2], puis à les convertir en matrices de cohérence [T3]. Les équations des matrices [S2] et [T3] sont présentées à la section 2.2.1.

La deuxième étape est la décomposition polarimétrique de la cible et est réalisée sous Matlab®. Elle consiste à décomposer la matrice de covariance en trois mécanismes de diffusion: 1) de surface (P_S), 2) de volume (P_V) et 3) de double réflexion (P_{DB}) selon l'équation (2.8) de la section 2.2.2.

Le modèle utilisé pour la diffusion de surface est le modèle de Bragg étendu (*X-Bragg*) proposé par Hajnsek *et al.* [2000], contrairement au modèle de Bragg utilisé par Freeman-Durden tel que décrit à la section 2.2.2. Ce modèle a été adopté parce qu'il a permis d'obtenir de meilleurs résultats sur des sols agricoles [Hajnsek *et al.*, 2009] et plus particulièrement pour des champs de maïs [Trudel *et al.*, 2012], qui sont nombreux dans la zone d'étude. La

rétrodiffusion de surface en polarisation HH ($\sigma^{\circ}_{\text{Surf_HH}}$) a aussi été calculée selon l'équation (2.10) de la section 2.2.2.

Finalement, les cinq nouvelles images obtenues (P_S , P_V , P_{DB} et $\sigma^{\circ}_{\text{Surf_HH}}$) ont été ajoutées comme de nouvelles bandes aux images brutes à l'aide de PCI Geomatica.

4.1.4 Orthorectification et correction radiométrique

Une procédure d'orthorectification utilisant un modèle numérique d'élévation (MNE) ainsi que des points de contrôle au sol a été réalisée sur l'ensemble des bandes à l'aide de PCI Geomatica. La fonction *Rational function Radar Satellite Modelling math model* a été utilisée. Dans le but de réduire le chatoiement dans les images de σ° , le filtre de Lee amélioré [Lopes *et al.*, 1990] avec une fenêtre de 7 x 7 pixels a été appliqué. Ce filtre a été privilégié pour sa capacité à préserver la texture tout en réduisant le chatoiement [*Ibid.*]. La taille de la fenêtre a été fixée suite à l'analyse visuelle des résultats obtenus avec diverses tailles de fenêtre. Compte tenu qu'un filtre moyenneur 7 x 7 est appliqué lors de la décomposition de la cible, la fenêtre de filtrage a été réduite au minimum alloué (3 x 3) pour les bandes P_S , P_V , P_{DB} et $\sigma^{\circ}_{\text{Surf_HH}}$. L'ensemble des images a été ré-échantillonné à une résolution spatiale de 12,5 m.

4.2 Traitement des données auxiliaires

Afin de prendre en compte différentes variables influençant σ° ainsi que les caractéristiques du sol et du site influençant l'écoulement de l'eau, cinq images auxiliaires ont été créées pour chacune des images RADARSAT-2 acquises.

Les données des pédo-paysages du Canada diffusées par le Service d'information sur les sols du Canada (SISCAN) ont été utilisées pour dériver trois images de texture : 1) % de sable, 2) % de limon et 3) % d'argile.

4) Les 27 images d'indice foliaire (ISF) composites de huit jours du capteur MODIS acquises ont été filtrées afin de ne conserver que les pixels avec des valeurs de ISF valides, soit entre 0 et 100 avec un facteur de mise à l'échelle de 0,01. Une image ISF a été associée à chacune des images RADARSAT-2 en fonction de sa date d'acquisition.

5) Les 30 jeux de données numériques d'élévation du Canada (DNEC) à l'échelle 1 : 50 000 ayant une précision altimétrique de 3 m ont été utilisés pour générer un MNE couvrant

l'ensemble de la zone d'étude. Le pourcentage de pente a ensuite été extrait afin de produire des images de pente.

Chaque groupe d'images a été ré-échantillonné à une résolution spatiale de 12,5 m et découpé afin d'obtenir des images ayant une empreinte au sol identique à l'image RADARSAT-2 correspondante et de manière à ce que les pixels se superposent. Ces images ont ensuite été ajoutées comme de nouvelles bandes aux images multibandes contenant les images orthorectifiées. L'ensemble des étapes de cette section a été réalisé à l'aide du système d'information géographique ArcGIS 10.1 [ESRI, 2011].

4.3 Création des polygones associés aux sondes

Pour chaque sonde d'humidité du sol installée, un polygone couvrant une superficie minimale de 62 500 m² a été numérisé dans un fichier de forme à partir d'une mosaïque d'orthophotos à l'aide d'ArcGIS 10.1 [ESRI, 2011]. Les orthophotos constituant la mosaïque possèdent une résolution spatiale de 50 cm et ont été acquises par le Ministère des Ressources Naturelles de l'Ontario en 2008. Le fichier de forme a ensuite été utilisé pour en générer 47 autres ne recouvrant que l'empreinte au sol de chacune des images RADARSAT-2.

Les polygones ont été délimités de manière à couvrir la plus grande zone hydrologiquement homogène possible. Ainsi, chaque polygone couvre une zone où la végétation et la pente sont uniformes afin qu'ils puissent être considérés représentatifs des conditions d'humidité du sol mesurées par la sonde correspondante. De plus, tel que mentionné à la section 3.3.1, l'emplacement des sondes est basé sur une analyse incluant la topographie, le type de sol et l'occupation du sol afin d'assurer une certaine homogénéité de l'humidité du sol dans la zone environnante. Bien que l'humidité du sol possède une variabilité qui peut être importante à l'intérieur d'un même champ, les contraintes opérationnelles telles que le nombre de champs suivis, la présence permanente des sondes et la culture des champs ont rendu nécessaires certains compromis.

En ce qui a trait à la superficie, 62 500 m² représente 400 pixels sur les images RADARSAT-2, ce qui est un nombre suffisamment grand pour éviter l'effet du chatoiement sur l'estimation du coefficient de rétrodiffusion et ainsi obtenir des résultats significatifs lors de calculs statistiques.

4.4 Création des intrants des réseaux de neurones

Pour chaque image, le nombre de pixels et la moyenne ont été calculés pour le polygone associé à chacune des sondes d'humidité présentes dans l'image. Les bandes traitées sont : l'angle d'incidence, σ_{HH}° , σ_{HV}° , P_S , P_V , P_{DB} , $\sigma_{Surf_HH}^\circ$, la pente, les proportions de sable, de limon et d'argile ainsi que l'ISF.

À partir des valeurs moyennes de σ° en polarisation HH, HV et VV, l'indice de végétation radar (IVR) a ensuite été calculé afin de permettre la prise en compte de l'effet de la végétation sur la rétrodiffusion radar. Bien que l'indice développé par Kim et van Zyl [2004] nécessite trois polarisations selon l'équation (4.3, Charbonneau *et al.* [2005] ont appliqué l'IVR lorsque seulement une polarisation parallèle est disponible en supposant que $\sigma_{HH}^\circ \approx \sigma_{VV}^\circ$, ce qui permet d'obtenir des IVR en double polarisation HH (IVR_{HH}) selon l'équation (4.4. L'IVR_{HH} en champs agricoles est toutefois légèrement plus faible que l'IVR lorsque la diffusion de surface est inférieure à 50 % [Trudel *et al.*, 2012].

$$IVR = \frac{8\sigma_{HV}^\circ}{\sigma_{HH}^\circ \sigma_{VV}^\circ + 2\sigma_{HV}^\circ} \quad (4.3)$$

$$IVR_{HH} = \frac{4\sigma_{HV}^\circ}{\sigma_{HH}^\circ + \sigma_{HV}^\circ} \quad (4.4)$$

Les 47 images RADARSAT-2 acquises couplées aux données auxiliaires, à l'IVR et à l'humidité du sol observée ont permis d'obtenir un jeu de données contenant 534 tuples contenant la polarisation HH, dont 208 contenant des données en polarisation HH et HV et 119 en polarisation quadruple sur lesquelles la décomposition polarimétrique a été possible. Par contre, seuls ces 119 tuples sont utilisés afin de pouvoir comparer les résultats obtenus avec les deux approches étudiées. Les histogrammes de fréquence des différentes variables pour ces 119 tuples sont présentés à la figure 4.4 et à la figure 4.5. Les histogrammes de σ° , de puissance et de σ° de surface sont distribués selon une distribution normale (figure 4.4), de même que les humidités du sol observées (figure 4.5b). L'indice IVR (figure 4.5c) est également distribué selon une distribution normale, mise à part une donnée aberrante. La distribution de la texture du sol (figure 4.5a) et de l'angle d'incidence (figure 4.5b) comporte des trous, qui peuvent affecter les performances des RNA.

Les jeux de données de 534 et 208 tuples ont pour leur part été utilisés afin d'évaluer l'impact de la taille de l'échantillon sur les performances des réseaux de neurones utilisant uniquement les données RADARSAT-2 en polarisation HH.

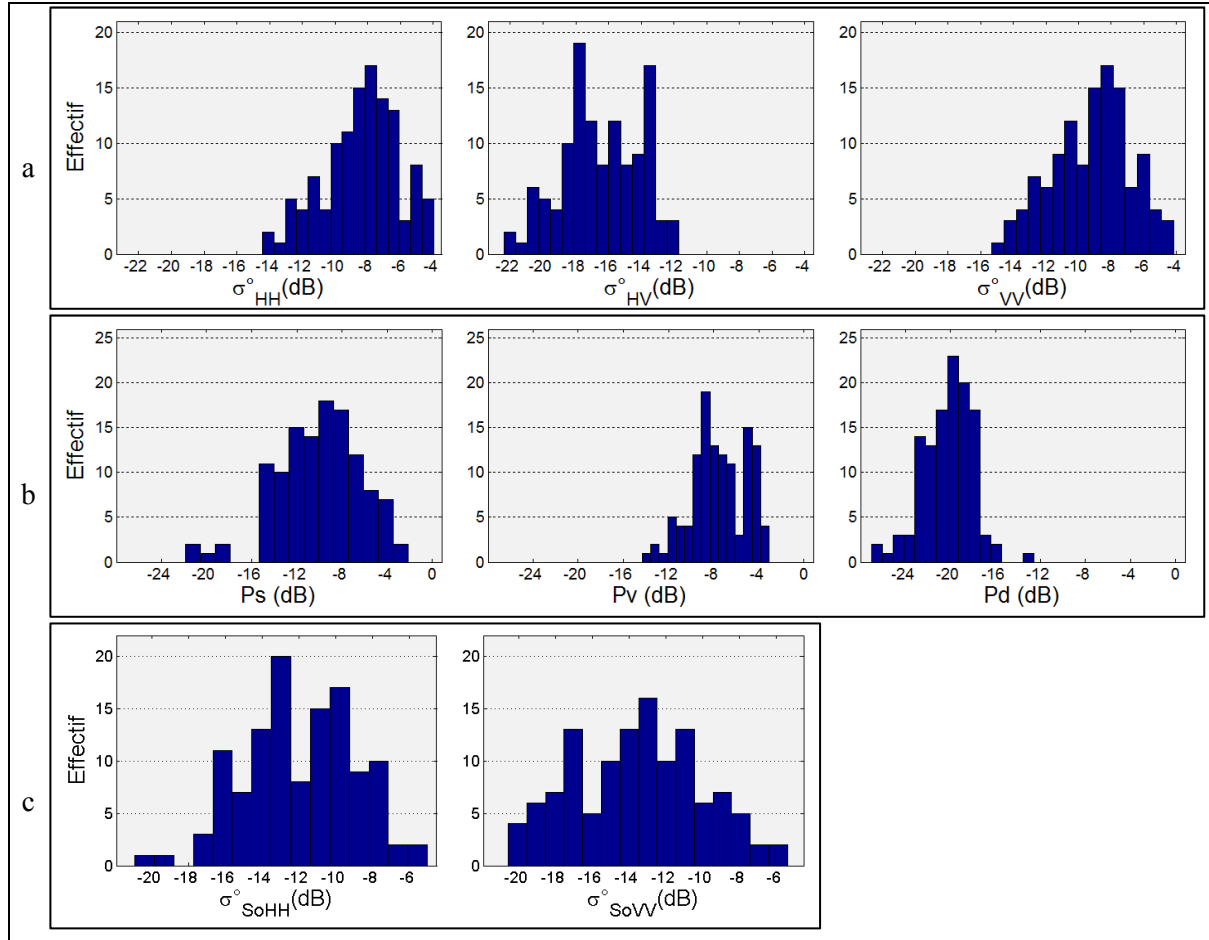


Figure 4.4 Histogrammes de fréquence de σ° (a), de puissance (b) et de σ° de surface (c)

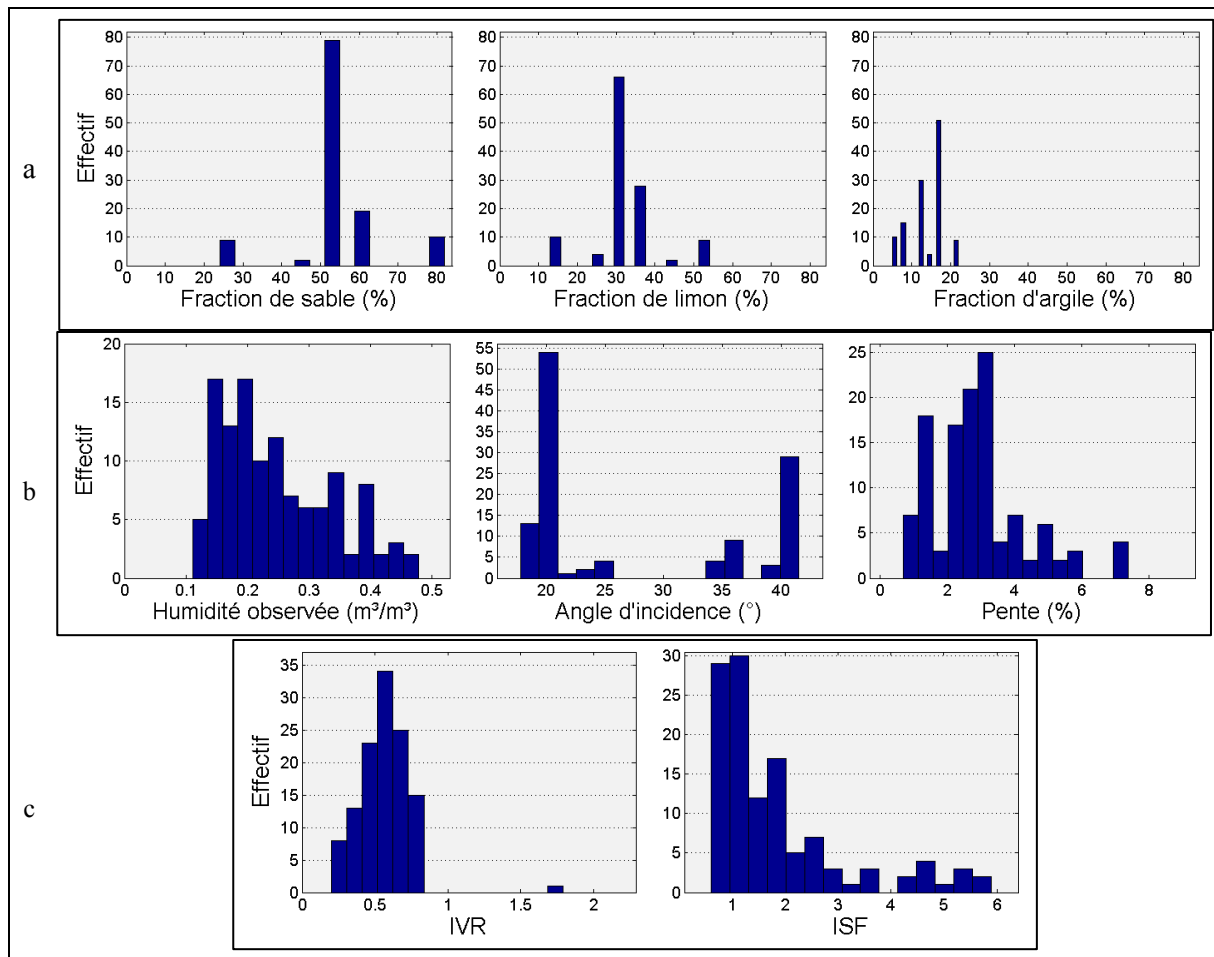


Figure 4.5 Histogrammes de fréquence de texture (a), d'humidité estimée, d'angle d'incidence et de pente (b), d'IVR et ISF (c)

4.5 Développement des réseaux de neurones artificiels (RNA)

L'ensemble des étapes liées aux réseaux de neurones artificiels a été réalisé à l'aide d'un script Matlab® utilisant les fonctions de la boîte à outils pour réseaux de neurones [The MathWorks, Inc., 2014]

4.5.1 Sélection des variables entrantes aux RNA

Compte tenu que les variables influençant l'humidité du sol et le coefficient de rétrodiffusion radar sont connues, la sélection a été basée sur les connaissances *a priori*. Les deux critères suivants ont été utilisés: 1) la sensibilité de la variable de sortie (i.e. l'humidité du sol) à cette variable et 2) sa disponibilité, ou la possibilité de la générer, en format matriciel. Ce dernier critère est nécessaire puisque l'un des objectifs de cette recherche est d'appliquer les réseaux

de neurones obtenus aux images RSO afin de générer des cartes d'humidité du sol et d'incertitude. Sur la base de ces deux critères, 15 variables ont été identifiées:

- les coefficients de rétrodiffusion radar en polarisation HH (σ°_{HH}) vu leur sensibilité à l'humidité du sol pour les valeurs de m_v inférieures à $0,35 \text{ m}^3/\text{m}^3$ [Dobson et Ulaby, 1981; Ulaby *et al.*, 1986];
- les fractions de sable et d'argile afin de prendre en compte l'effet de la texture sur la constante diélectrique [Schmugge, 1983] et sur le taux d'infiltration;
- la pente en raison de son effet sur les taux d'infiltration et de ruissellement;
- l'angle d'incidence (θ°) pour son influence sur la magnitude de la rétrodiffusion en fonction de la rugosité de surface et de la végétation [Ulaby et Batlivala, 1976];
- le coefficient de rétrodiffusion radar en polarisation HV (σ°_{HV}) parce qu'il permet d'éliminer pratiquement la sensibilité du signal radar à la direction des rangs de culture relativement à celle de visée [McNairn et Brisco, 2004], et en raison de sa sensibilité accrue à l'effet dépolarisant de la végétation et moindre à l'humidité du sol par rapport à σ°_{HH} [Romshoo *et al.*, 2002];
- l'indice de surface foliaire (ISF) parce qu'il est un bon indicateur de la croissance et de la densité de la biomasse végétale ainsi que de la densité d'un peuplement [Weiss, 1998]. Son domaine de valeurs est $[0, 10]$ mais n'atteint généralement pas plus de 8 pour une forêt tempérée dense, de 4,3 pour des cultures de maïs et de 6,4 si irriguées, alors qu'il est de 3,2 et 5,6 pour le soya [Vina, 2004];
- l'indice de végétation radar (IVR) pour sa capacité à caractériser la quantité de végétation présente ainsi que la diffusion de surface [Kim et van Zyl, 2009];
- la puissance de diffusion de volume (P_v) pour caractériser la diffusion de volume associée à la végétation et l'atténuation de la rétrodiffusion qu'elle cause [Attema et Ulaby, 1978; Freeman et Durden, 1998];
- le coefficient de rétrodiffusion en polarisation HH provenant uniquement de la surface ($\sigma^{\circ}_{\text{Surf_HH}}$) et la puissance de diffusion de surface (P_s) pour améliorer la modélisation de la diffusion de surface [Freeman et Durden, 1998].

4.5.2 Architecture des réseaux

Les réseaux de neurones développés sont des perceptrons multicouches (PMC) possédant une architecture variable en fonction du nombre d'intrants et de la taille de la couche cachée. Ce type de RNA possède une architecture non-bouclée et des connexions unidirectionnelles vers l'avant [*feed-forward*] complètes entre les différentes couches. Ce qui signifie que tous les neurones d'une couche sont connectés avec chacun des neurones de la couche suivante et que l'information s'y propage vers l'avant, soit à partir de chacun des neurones de la couche d'entrée vers l'ensemble de ceux de la (ou des) couche(s) cachée(s), puis au neurone de la couche de sortie. La propagation de l'information entre les couches est contrôlée par la fonction d'activation du neurone qui transforme l'information provenant de la couche précédente à l'aide d'une fonction mathématique.

Tous les réseaux développés dans cette recherche comprennent trois couches : une couche d'entrée, une cachée et une de sortie. Une seule couche cachée a été utilisée puisqu'il a été démontré que les PMC à une couche cachée constituée d'un nombre suffisant de neurones utilisant une fonction non-linéaire sont capables d'estimer n'importe quelle fonction continue de nombres réels [Dawson *et al.*, 1997]. Plus spécifiquement, plusieurs recherches portant sur l'inversion de l'humidité du sol à l'aide de données RSO ont utilisé des RNA à une couche cachée avec succès [Baghdadi *et al.*, 2002a; Baghdadi *et al.*, 2016; Pierdicca *et al.*, 2008; Satalino *et al.*, 2002]. La couche d'entrée contient un nombre variable de neurones représentant chacun une variable alors que la couche de sortie ne contient qu'un seul neurone correspondant à m_v . Ces deux couches sont reliées par la couche dite cachée contenant un nombre variable de neurones. Un terme de biais est aussi présent pour la couche cachée et celle de sortie. Les biais sont représentés par un neurone additionnel dans la couche précédente et la fonction d'activation est toujours égale à 1. La figure 4.6 présente une schématisation de la structure utilisée.

La fonction d'activation sigmoïdale tangente hyperbolique a été choisie pour les neurones de la couche cachée. Ce choix est basé sur sa capacité d'ajuster les poids de connexion plus efficacement que les fonctions sigmoïdales logistique et linéaire, ce qui permet un apprentissage plus rapide et produit des résultats légèrement supérieurs par rapport à la fonction sigmoïdale logistique [Harrington, 1993; Maier et Dandy, 1998]. D'autre part, la

fonction identité, qui n'effectue aucune transformation de l'information, a été utilisée pour le neurone de la couche de sortie.

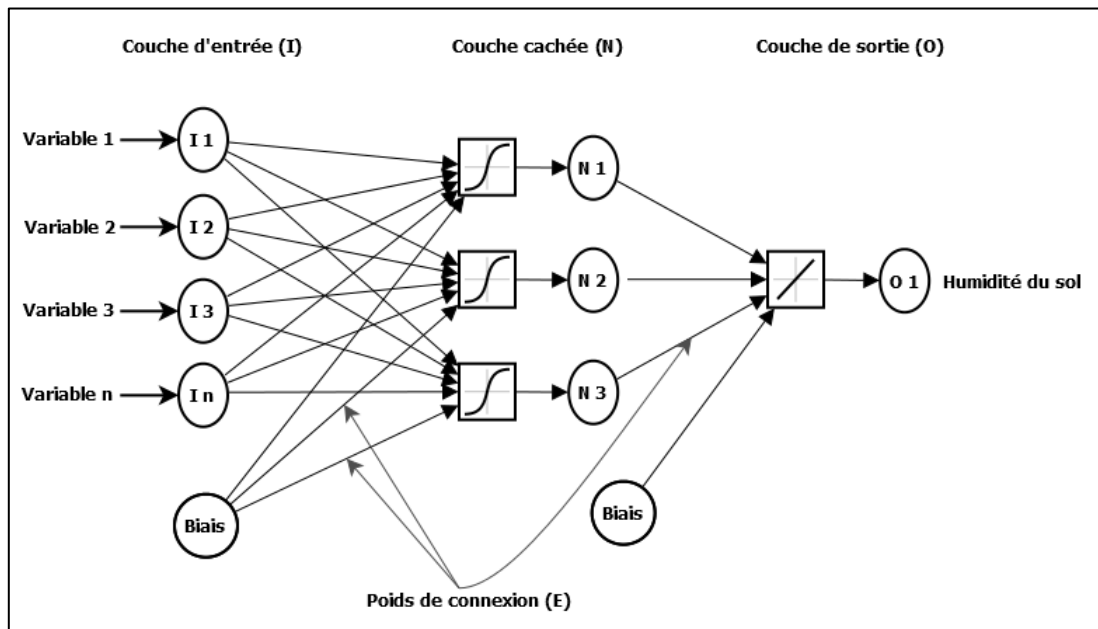


Figure 4.6 Structure des réseaux de neurones développés

Plusieurs auteurs ont publié des lignes directrices afin de déterminer le nombre optimal de neurones à inclure dans la couche cachée. Plusieurs de celles-ci utilisent un ratio entre la taille du réseau et la taille de l'échantillon d'entraînement. Toutefois, ces lignes directrices ne garantissent pas une géométrie optimale. Ainsi, traditionnellement la géométrie est établie par essai - erreur [Maier et Dandy, 2000]. Par conséquent, des réseaux avec une couche cachée contenant entre 1 et 20 neurones ont été entraînés pour chacune des combinaisons de variables entrantes évaluées.

Les intrants sélectionnés à la section 4.5.1 ont été classés en quatre groupes d'intrants. Le premier groupe comprend la rétrodiffusion radar copolarisée horizontalement (σ°_{HH}) ainsi que l'angle d'incidence. Le deuxième groupe inclut les intrants relatifs à la topographie et à la texture du sol, soit le pourcentage de pente ainsi que les fractions de sable et d'argile. La fraction de limon n'a pas été utilisée puisque qu'elle n'apporte pas d'information supplémentaire car la somme des fractions égale toujours 100 %. De plus, son utilisation augmenterait la redondance dans les données entrantes et ralentirait le processus d'apprentissage [Rojas, 1996]. Le troisième groupe comprend les intrants relatifs à la

végétation ou permettant de la prendre en compte. Il est constitué des quatre variables suivantes : 1) l'indice de surface foliaire (ISF), 2) l'indice de végétation radar (IVR), 3) les données radar en polarisation croisée (σ°_{HV}) et 4) la puissance de diffusion de volume (P_v). Le quatrième groupe contient les intrants relatifs à la rétrodiffusion de surface, soit la puissance de diffusion de surface (P_s) ainsi que la rétrodiffusion radar de surface copolarisée horizontalement ($\sigma^{\circ}_{Surf_HH}$).

En considérant les variables entrantes évaluées, la géométrie et le degré de connectivité utilisés, 62 combinaisons de variables différentes sont possibles. L'architecture des RNA analysés a été établie de manière progressive plutôt que par l'analyse des 62 combinaisons. Plus précisément, les intrants du premier groupe ont été utilisés pour créer le RNA de base pour l'approche *Polarisation simple et double*. Les intrants des groupes 2 et 3 (à l'exception de P_v) ont été ajoutés individuellement de manière successive et les résultats obtenus analysés. Lorsque l'ajout d'un intrant augmentait les performances et que la distribution spatiale de l'humidité du sol estimée était plausible, l'intrant a été conservé dans l'architecture finale de l'approche *Polarisation simple et double*. Pour l'approche *Polarimétrique*, les intrants du groupe 4 ainsi que P_v ont été utilisés successivement pour remplacer leurs équivalents provenant des données en polarisation simple et multiple. En particulier, la rétrodiffusion de surface en polarisation HH ($\sigma^{\circ}_{Surf_HH}$) et la puissance de diffusion de surface (P_s) ont été employées en lieu de σ°_{HH} ; la puissance de diffusion de volume (P_v) a été utilisée en lieu de σ°_{HV} , de l'ISF et de l'IVR. L'angle d'incidence est la seule variable présente dans l'ensemble des architectures testées.

L'architecture des RNA utilisés pour l'évaluation de l'influence de la taille de l'échantillon sur les performances a été conditionnée par la meilleure combinaison d'intrants obtenue à partir des intrants provenant des groupes 1 et 2 puisqu'ils sont les seuls disponibles pour l'ensemble des 534 tuples obtenus.

4.5.3 Entraînement, validation et test des RNA

Une fois les variables entrantes sélectionnées et l'architecture définie, la création proprement dite des réseaux de neurones a été réalisée en trois phases distinctes: l'entraînement, la validation et le test. Le jeu de données initial a par conséquent été divisé aléatoirement en trois échantillons, soit: 1) entraînement, 2) validation et 3) test, contenant respectivement 70 %, 15

% et 15 % du jeu de données. Plusieurs auteurs ont proposé divers pourcentages pour chacun des échantillons variant de 60% à 65% pour l'entraînement, de 10% à 30 % pour la validation et de 20% à 25 % pour le test [Looney, 1996; Nelson et Illingworth, 1991; Rodriguez-Fernandez *et al.*, 2015; Swingler, 2001] et de manière aléatoire, notamment par Rodriguez-Fernandez *et al.* [2015]. Compte tenu de la petite taille de notre jeu de données, un pourcentage plus élevé a été attribué à l'entraînement afin de maximiser la capacité d'apprentissage des RNA. Les données restantes ont été divisées également entre les échantillons de validation et de test afin d'être en mesure d'effectuer une validation suffisante et d'évaluer sa capacité de généralisation lors de la phase de test. Ces deux éléments sont abordés plus en détail ultérieurement.

Il peut être intéressant de faire un parallèle entre la modélisation hydrologique et les RNA. Le calage d'un modèle hydrologique correspond à la phase d'entraînement du RNA et la validation du modèle hydrologique à la phase de test du RNA. Le développement d'un RNA comprend cependant une phase de validation supplémentaire, au cours de laquelle des données indépendantes sont utilisées pour évaluer les performances du RNA durant la phase d'entraînement afin d'y mettre fin lorsque certaines conditions sont respectées.

Les réseaux de neurones utilisés dans cette recherche ont été entraînés avec l'algorithme Levenberg-Marquardt (LM) [Levenberg, 1944; Marquardt, 1963]. Cet algorithme a été utilisé dans plusieurs études portant sur l'humidité du sol [Kseneman *et al.*, 2012; Notarnicola *et al.*, 2008; Rodriguez-Fernandez *et al.*, 2015]. Il s'agit d'un algorithme d'apprentissage supervisé à réglage avant avec rétro-propagation de l'erreur, désigné dans la littérature selon le terme *feedforward error-backpropagation*. Comme il s'agit d'un apprentissage supervisé, chaque échantillon utilisé contient les valeurs des différentes variables entrantes utilisées ainsi que la valeur de sortie (m_v) correspondante. Avant l'entraînement, les valeurs de chacune des variables sont préalablement normalisées afin d'obtenir un domaine de valeurs [-1 1], identique à celui de la fonction d'activation. Les poids de connexion des neurones sont initialisés aléatoirement mais normalisés à l'aide de l'algorithme Nguyen-Widrow [Nguyen et Widrow, 1990].

Le processus d'entraînement se déroule en plusieurs itérations qui se composent de deux étapes: 1) le réglage avant et 2) la rétro-propagation de l'erreur. Les réseaux de neurones

entraînés avec l'algorithme LM procèdent à un apprentissage par lots (*batch training*), ce qui signifie qu'une itération est de la même taille que l'échantillon d'apprentissage. L'ajustement des poids de connexion a donc été réalisé après que tous les tuples du jeu de données d'entraînement aient été présentés au RNA.

Au cours du réglage avant, tous les tuples composant le jeu de données d'entraînement sont présentés à la couche d'entrée l'un après l'autre. Pour chacun d'eux, les sommes pondérées des différents neurones d'entrée sont traitées par chacun des neurones de la couche cachée par la fonction de transfert, avant que l'information ne passe à la couche de sortie. L'erreur obtenue pour chaque tuple est stockée dans un vecteur. La fonction-objectif utilisée pour évaluer les performances du réseau est l'erreur quadratique moyenne (MSE) sur les estimations de m_v .

Lors de la rétro-propagation de l'erreur, l'erreur quadratique moyenne calculée à partir du vecteur d'erreur a été utilisée pour ajuster les poids de connexion afin de la minimiser. Pour l'algorithme LM, cet ajustement est géré par trois paramètres liés au momentum (μ): le momentum initial et les facteurs d'augmentation et de réduction du momentum. Les valeurs par défaut de 0,001, 10 et 0,1 sont utilisées. Le momentum (μ) correspond à la proportion de la variation de poids précédente ajoutée au changement de poids en cours [Maier et Dandy, 1998]. Ce paramètre permet de renforcer les tendances générales et d'atténuer le comportement oscillatoire de l'erreur lors de l'entraînement [NeuralWare, Inc., 1991]. Le facteur d'augmentation ou de réduction est appliqué à μ dans le cas où l'erreur est plus élevée ou plus faible que lors de l'itération précédente, respectivement. L'ordre de présentation des différents tuples d'apprentissage est modifié lors des itérations suivantes et la mise à jour des poids est effectuée à chacune d'elles.

La phase de validation est réalisée après chacune des itérations. Elle consiste à présenter l'échantillon de validation au RNA obtenu et à calculer l'erreur quadratique moyenne sur ces données. Cette phase permet de suivre l'évolution des performances du RNA et d'évaluer sa capacité de généralisation.

L'entraînement se poursuit jusqu'à ce que l'une des conditions d'arrêt suivantes soit rencontrée:

- le nombre maximal de 50 itérations est atteint;

- la performance atteint l'objectif de $MSE = 0,0$;
- le gradient de performance minimal de $1e-7$ est atteint;
- le momentum maximal $1e10$ est atteint;
- la performance en validation se détériore durant 6 itérations consécutives.

Le nombre maximal d'itérations a été fixé à 50 dans le but d'atteindre un équilibre entre la précision des résultats obtenus et le temps de calcul nécessaire. De plus, si un nombre plus élevé d'itérations est nécessaire, cela indique que l'information contenue dans les données d'entraînement est insuffisante [Nemesio Rodriguez-Fernandez, Communication personnelle, 2014]. Tout comme plusieurs auteurs l'ont fait (*ex.* [Bowden *et al.*, 2002; Notarnicola *et al.*, 2008; Posa *et al.*, 2004]), les valeurs par défaut de gradient de performance minimum et de momentum maximal ont été utilisées. La dernière condition est liée à l'utilisation de la méthode d'arrêt précoce qui permet de minimiser le risque de surentraînement sans toutefois arrêter l'entraînement prématurément [Amari *et al.*, 1997].

La phase de test est ensuite réalisée une fois l'entraînement terminé. Le RNA est testé avec les données de l'échantillon de test afin d'évaluer sa capacité de généralisation. Les résultats obtenus n'ont toutefois pas d'influence sur l'entraînement et sont principalement utiles pour comparer différents modèles entre eux.

Afin de maximiser la probabilité d'obtenir les meilleurs réseaux de neurones possible avec les données disponibles, 2000 essais distincts ont été réalisés pour chacune des architectures testées. Ce nombre a été établi suite à des tests qui ont démontré que le fait d'effectuer davantage d'essais ne permettait pas d'obtenir des RNA avec de meilleures performances, et donc, que des solutions optimales avaient été trouvées.

Le tableau 4.1 présente un sommaire des paramètres utilisés lors des phases d'entraînement, de validation et de test.

Tableau 4.1 Sommaire des paramètres utilisés pour l'entraînement, la validation et le test

Paramètre	Valeur	Paramètre	Valeur
Algorithme d'entraînement	Levenberg-Marquardt	Facteur d'augmentation du momentum	10
Algorithme de division des données	Aléatoire	Facteur de réduction du momentum	0,1
Ratio de l'échantillon d'entraînement	70%	Momentum maximum	1 e10
Ratio de l'échantillon de validation	15%	Gradient de performance minimum	1 e-7
Ratio de l'échantillon de test	15%	Fonction de transfert	Tansig
Fonction de performance	MSE	Nombre maximal d'itérations	50
Objectif de performance	0,0	Nombre maximal de validations échouées	6
Momentum initial	0,001	Nombre de cycles d'entraînement	2000

4.5.4 élection du critère de performance

La performance des réseaux de neurones peut être évaluée selon différents critères, principalement la vitesse d'entraînement, le temps d'exécution du modèle et la précision des prévisions. Comme l'objectif est d'obtenir les résultats les plus près de la réalité possible, c'est ce dernier critère qui a été utilisé pour la sélection finale des réseaux qui serviront à évaluer chacune des deux approches étudiées.

4.5.5 Sélection des RNA

Suite à l'exécution des 2000 essais pour chacune des combinaisons testées, un total de plus de 2200 RNA a été obtenu et une étape de sélection a été réalisée. Afin d'évaluer l'incertitude reliée au réseau de neurones, les 30 meilleurs RNA sont identifiés pour chaque combinaison d'intrants et seront utilisés pour analyser la variabilité des RNA. La sélection des RNA a été réalisée en trois étapes successives. L'objectif est de ne conserver que ceux 1) qui ont été obtenus avec des échantillons représentatifs du jeu de données complet, 2) pour lesquels les coefficients de corrélation obtenus pour les échantillons d'entraînement et de test entre les valeurs de m_v estimées (m_{ve}) et observées (m_{vo}) ne sont pas significativement différents et 3) qui offrent les meilleures performances.

La première étape vise à éliminer les RNA obtenus à partir d'échantillons non-représentatifs du jeu de données complet, ce qui peut se traduire par une différence marquée entre les performances obtenues avec les échantillons d'entraînement et de validation [Maier et Dandy, 2000]. La représentativité des échantillons a été vérifiée au niveau de la moyenne et de la variance à l'aide des tests de Student et de Fisher, équations (4.5 et (4.6) respectivement, avec un niveau de confiance de 95%. Ainsi, les RNA pour lesquels l'une ou l'autre des *p-value* obtenues est inférieure à 0,05 sont exclus.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad (4.5)$$

$$F = \frac{S_1^2}{S_2^2} \quad (4.6)$$

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} \quad (4.7)$$

où \bar{x}_1 et \bar{x}_2 sont les moyennes de l'échantillon et du jeu de données complet, S_1 et S_2 les écarts-types et n_1 et n_2 , la taille de l'échantillon et du jeu de données complet, respectivement.

La deuxième étape vise à identifier les RNA candidats. Les coefficients de corrélation de Pearson (r) pour les échantillons d'entraînement et de test sont calculés et analysés afin de déterminer si l'écart est significatif. Les coefficients de Pearson ont d'abord été convertis en des variables z distribuées normalement (*z-score*) à l'aide de la transformation de Fisher (équation (4.8) à partir desquels la valeur Z (*Z-value*) a été obtenue (équation (4.9) puis associée à sa *p-value*. Les RNA ont été classés comme des réseaux candidats lorsque l'écart est non-significatif avec un intervalle de confiance de 95% (*p-value* > 0,05).

$$z(r) = 0,5 \ln \left(\frac{(1+r)}{(1-r)} \right) \quad (4.8)$$

$$Z = \frac{(z_1 - z_2)}{\sqrt{\left(\frac{1}{n_1 - 3} \right) + \left(\frac{1}{n_2 - 3} \right)}} \quad (4.9)$$

où r est le coefficient de Pearson, z_1 et z_2 sont les z -score des échantillons d'entraînement et de test et n_1 , n_2 la taille des échantillons d'entraînement et de test.

La troisième étape consiste à sélectionner les 30 meilleurs RNA pour chacune des combinaisons testées en fonction des performances des réseaux de neurones. Celles-ci sont évaluées à l'aide des quatre critères suivants : 1) le coefficient de corrélation de Pearson (r), 2) l'erreur-type (RMSE), 3) l'erreur absolue moyenne (MAE) et 4) l'erreur relative moyenne en pourcent (MRE). Ces critères de performance sont calculés à l'aide des équations (4.10) à (4.13).

Le meilleur des 30 meilleurs RNA de chacune des combinaisons a été utilisé pour : 1) l'analyse statistique des performances; et 2) comparer les performances des différentes combinaisons évaluées. Les 30 RNA ont pour leur part été utilisés pour produire les cartes d'estimation de l'humidité du sol et son incertitude. Celles-ci sont utilisées pour : 1) l'analyse de la répartition spatiale de l'humidité du sol estimée et de l'incertitude d'un point de vue visuel sur l'ensemble du territoire couvert par les cartes; et 2) l'analyse statistique pour deux sous-bassins versants.

$$r = \left[\frac{\sum_{i=1}^N (y_i^m - \overline{y_i^m})(y_i^s - \overline{y_i^s})}{\sqrt{\sum_{i=1}^N (y_i^m - \overline{y_i^m})^2 \sum_{i=1}^N (y_i^s - \overline{y_i^s})^2}} \right] \quad (4.10)$$

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^N (y_i^m - y_i^s)^2 \right]^{0.5} \quad (4.11)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i^m - y_i^s| \quad (4.12)$$

$$MRE = \frac{100}{N} \sum_{i=1}^N \frac{|y_i^m - y_i^s|}{y_i^m} \quad (4.13)$$

où y_i^m et y_i^s sont les valeurs de m_v mesurées et simulées et $\overline{y_i^m}$ and $\overline{y_i^s}$ les valeurs moyennes de m_v mesurées et simulées, dans l'ordre et N la taille de l'échantillon.

4.6 Création des cartes d'humidité du sol estimée et d'incertitude

L'étape de production des cartes d'estimation de l'humidité du sol et de l'incertitude a été réalisée à l'aide de scripts Matlab®. Pour chaque combinaison d'intrants, les 30 RNA sélectionnés ont été exécutés pour chacune des images sélectionnées afin d'obtenir une matrice contenant les 30 valeurs d'humidité estimées pour chaque pixel. La moyenne et l'écart-type ont ensuite été calculés pour chaque cellule de la matrice afin d'obtenir les matrices d'humidité estimée moyenne et d'incertitude, dans l'ordre. Ces deux matrices ont finalement été enregistrées en format matriciel géoréférencé.

Les images utilisées pour analyser les RNA ont été sélectionnées afin de permettre l'évaluation de l'influence des intrants sur l'estimation de l'humidité du sol et son incertitude mais aussi la sensibilité des RNA à l'humidité du sol, à l'angle d'incidence et à la quantité de végétation présente. De manière à pouvoir atteindre ces objectifs, les images ont été sélectionnées en fonction des critères suivants : 1) tous les intrants utilisés sont disponibles; 2) elles présentent différentes conditions d'humidité du sol; 3) elles ont été acquises à des stades de développement végétatif variés; 4) elles ont été acquises à divers angles d'incidence; et 5) elles possèdent une zone de recouvrement la plus grande possible. Le premier critère fait en sorte que seules les images acquises en mode polarimétrique ont été considérées puisqu'elles sont les seules à permettre l'application des RNA associés à chacune des deux approches. Au final, des cartes d'humidité du sol estimée et d'incertitude ont été produites à partir de cinq images RADARSAT-2.

Les conditions d'humidité du sol lors de l'acquisition des images RADARSAT-2 ont été établies à partir de l'humidité du sol et des précipitations antécédentes mesurées sur le terrain. Ces données ont par la suite été comparées aux résultats afin d'évaluer la sensibilité des RNA à l'humidité du sol et à la végétation. Les données utilisées proviennent de la sonde 9915 et de la station météorologique Baldwin (voir la figure 3.2 pour l'emplacement des sites) puisqu'elles sont les plus centrales par rapport à la zone de recouvrement (voir la figure 4.8). La figure 4.7 présente les mesures d'humidité du sol et de précipitations à ces sites.

Les images du 1^{er} et du 25 mai 2012 ont été acquises avec une végétation peu ou pas développée lorsque les conditions de sols étaient humides ($0,43 \text{ m}^3/\text{m}^3$) et sèches ($0,20 \text{ m}^3/\text{m}^3$),

respectivement. Lorsque la végétation était bien développée, les images du 15 juillet 2012 et du 31 juillet 2013 ont été acquises dans des conditions de sol sec et relativement humide ($0,13 \text{ m}^3/\text{m}^3$ et $0,20 \text{ m}^3/\text{m}^3$, respectivement) et celle du 8 août 2012, dans des conditions relativement humides ($0,24 \text{ m}^3/\text{m}^3$). Les images du 15 juillet 2012 et du 31 juillet 2013 ont été acquises à des angles d'incidence de $37^\circ\text{--}39^\circ$ et $36^\circ\text{--}38^\circ$, respectivement, comparativement à $18^\circ\text{--}21^\circ$ pour les autres. Le tableau 4.2 présente le sommaire des images utilisées et les conditions lors de l'acquisition.

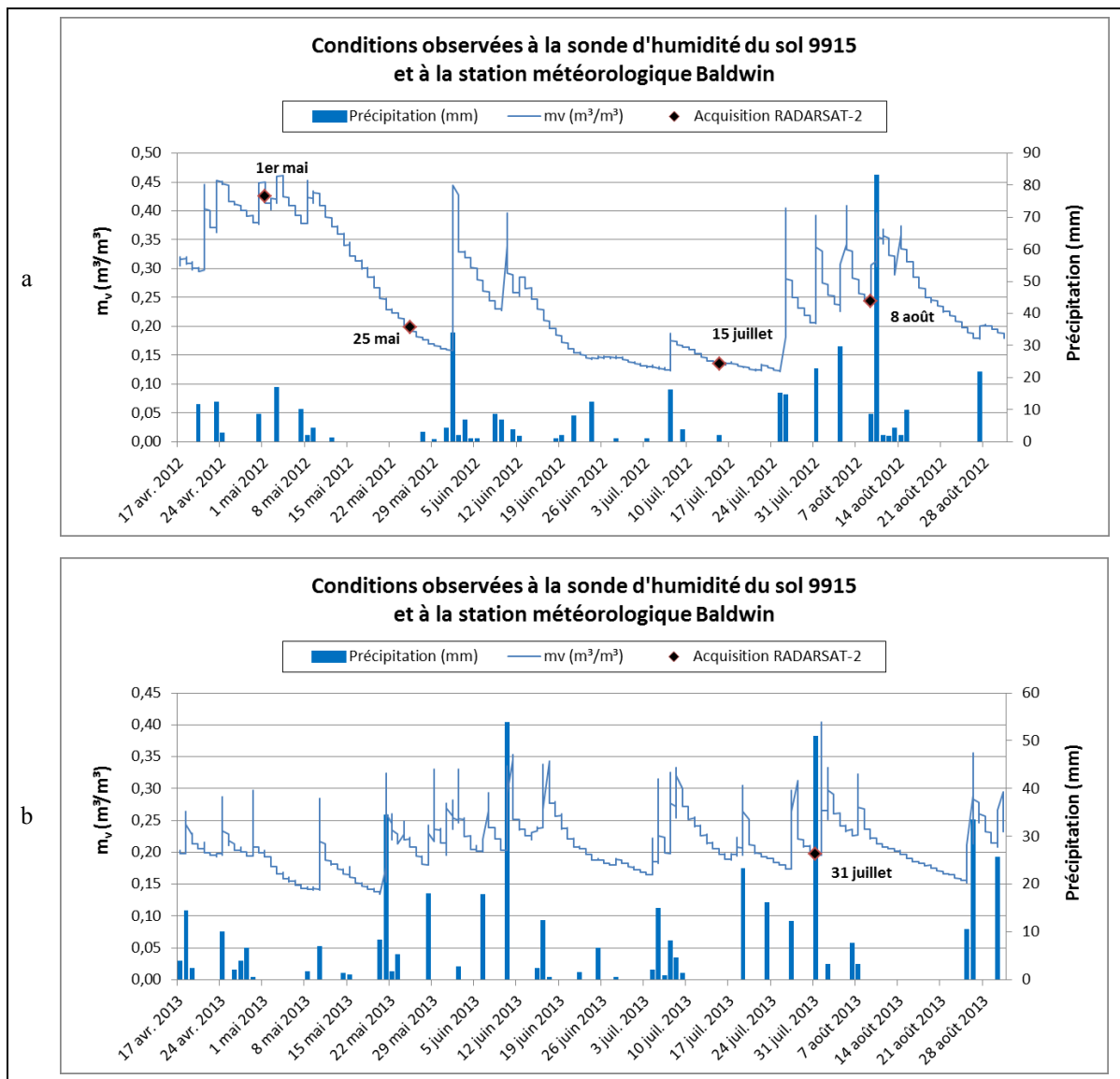


Figure 4.7 Humidité du sol et précipitations mesurées in situ en 2012 (a) et 2013 (b)

Tableau 4.2 Conditions observées lors de l'acquisition des images RADARSAT-2

Date d'acquisition de l'image RS-2	Angle d'incidence (°)	Humidité mesurée (m ³ /m ³)	Condition d'humidité du sol représentée	Stade de développement de la végétation
1 ^{er} mai 2012	18-21	0,43	Très humide	Peu ou pas développée
25 mai 2012	18-21	0,20	Relativement humide	Peu ou pas développée
15 juillet 2012	37-39	0,13	Sèche	Bien développée
31 juillet 2013	18-21	0,20	Relativement humide	Bien développée
8 août 2012	36-38	0,24	Humide	Bien développée

Il est toutefois à noter que malgré l'humidité observée les 25 mai 2012 et 31 juillet 2013 soit identique, 51 mm de pluie ont été enregistrés le 31 juillet 2013 alors qu'aucune précipitation significative n'a été enregistrée dans les semaines précédant le 25 mai 2012. Compte tenu que les données de précipitations sont journalières, on peut supposer que seulement une partie des précipitations avait été reçue lors de l'acquisition de l'image à 11h39, heure UTC. Il est également possible qu'il y ait eu des précipitations au moment de l'acquisition.

Les sous-bassins versants Baldwin et Pepperlaw ont été utilisés pour les analyses statistiques des cartes. Les empreintes au sol des images analysées, les sous-bassins versants, la localisation des sondes d'humidité ainsi que des stations météorologiques sont présentées à la figure 4.8.

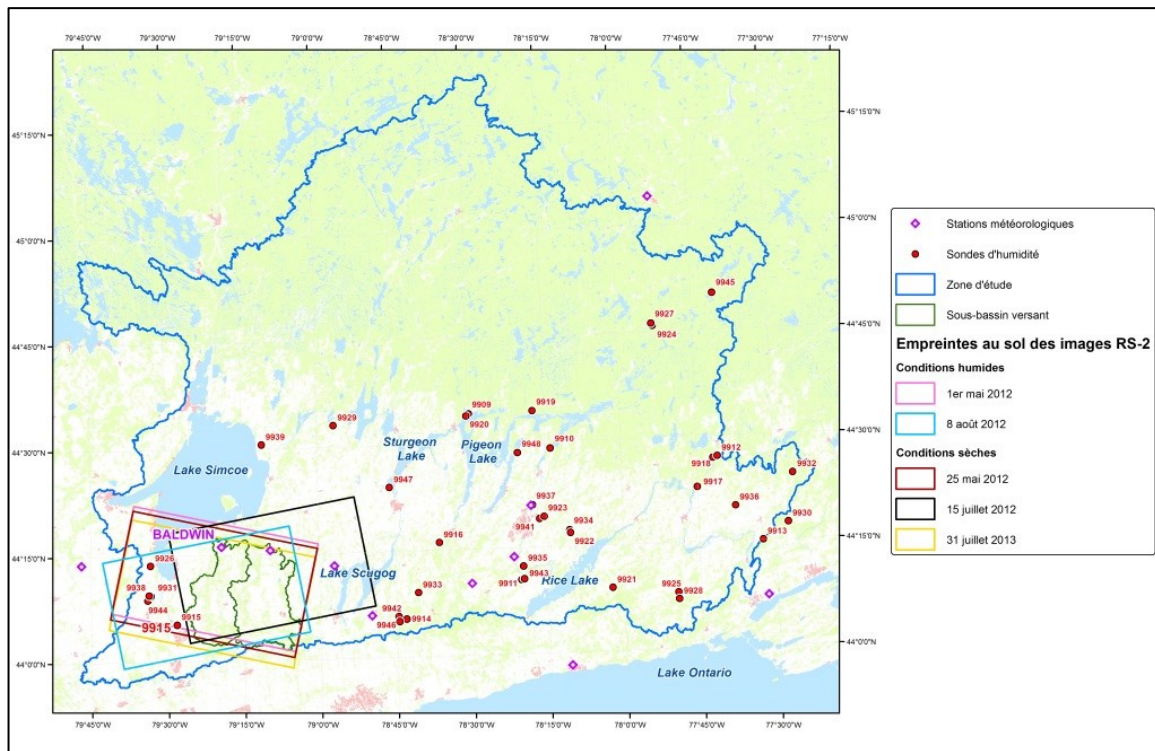


Figure 4.8 Carte de localisation des images utilisées pour la production des cartes d'estimation d'humidité ainsi que des sondes et stations météorologiques

La zone couverte par les cartes produites est principalement utilisée à des fins agricoles (en majorité du maïs), à l'exception des sols mal drainés, principalement occupés par des marécages, des tourbières et des sols à faible potentiel agricole où le couvert forestier a été conservé (figure 4.9). Les marécages se définissent comme des milieux humides où le substrat est généralement saturé en eau de manière continue et où la végétation est composée d'herbes, de mousses et d'une couverture minimale de 25 % d'arbres ou de grands arbustes et habituellement sans accumulation importante de tourbe. Les marécages sont fréquemment inondés au printemps et les accumulations y sont plus ou moins persistantes [Ontario Ministry of Natural Resources, 2008]. Les marais sont typiquement composés d'une mosaïque de végétation dominée par des plantes ligneuses émergentes, de canaux d'écoulement et de bassins d'eau libre où la végétation aquatique prolifère [*Ibid.*]. Une zone particulière d'une superficie de 7 360 ha située au sud du lac Simcoe, connue sous le nom des marais de Holland (*Holland Marsh*), a été drainée à 60 % afin de créer deux polders utilisés pour la culture maraîchère, à 70 % des carottes et des oignons en 2006 [Planscape, 2009]. Les pixels

représentant les forêts, les zones anthropiques imperméables ainsi que les voies de communication ont été filtrés et exclus des résultats et analyses.

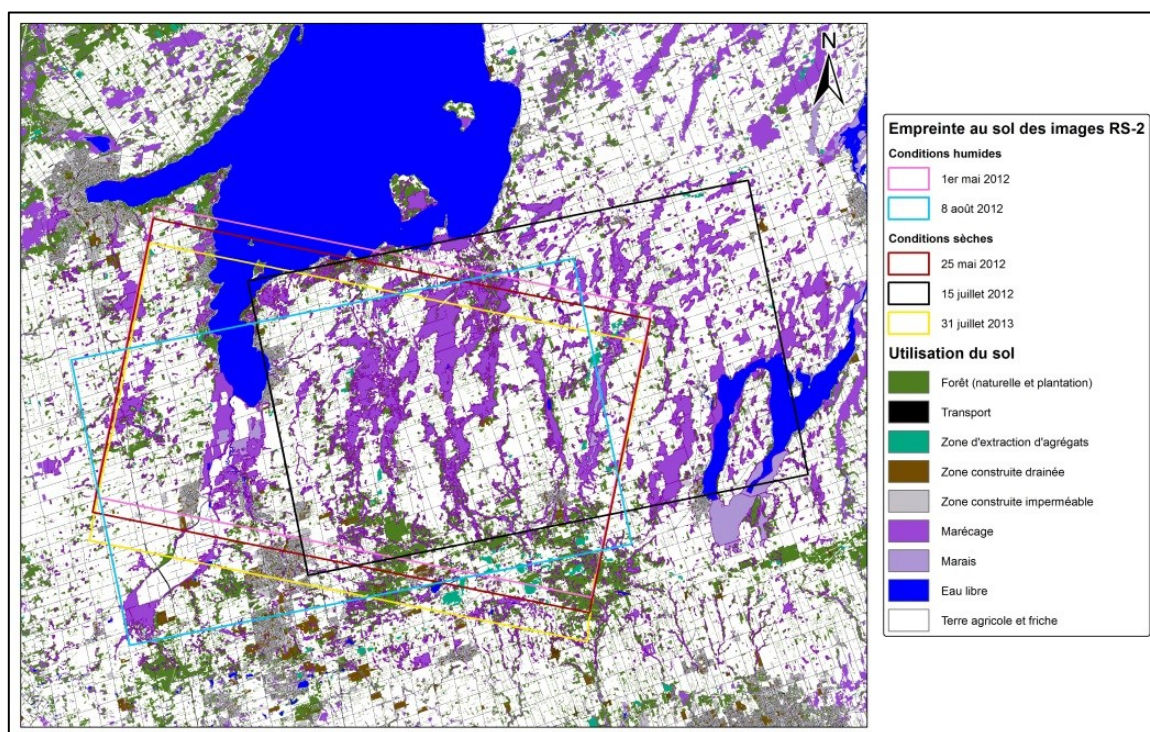


Figure 4.9 Carte de l'utilisation des sols dans la zone de recouvrement des images RADARSAT-2 (source: [Ontario Ministry of Natural Resources, 2008])

CHAPITRE 5 RÉSULTATS ET DISCUSSION

Ce chapitre est scindé en deux sections distinctes. La première partie correspond à l'approche *Polarisation simple et double* qui comprend les résultats obtenus à l'aide des données radar co-polarisées et en polarisation croisée, soit les images acquises en mode Large et Standard en polarisation HH et HH/HV, et l'analyse de ceux-ci. La seconde partie correspond à l'approche *Polarimétrique* et présente les résultats obtenus à l'aide de données radar polarimétriques, soit les images acquises en mode Standard Quad-Pol en polarisation HH, HV, VH et VV, et l'analyse de ceux-ci. Chaque section présente dans un premier temps les performances du meilleur RNA puis les cartes d'estimation de l'humidité du sol et de l'incertitude sur l'estimation.

5.1 Approche polarisation simple et double

5.1.1 Angle d'incidence et coefficient de rétrodiffusion HH

Le premier RNA développé utilise uniquement l'angle d'incidence (θ°) et le coefficient de rétrodiffusion co-polarisé horizontalement (σ°_{HH}) comme variables entrantes.

Le RNA le plus performant parmi les 30 RNA sélectionnés pour cette combinaison d'intrants possède une couche cachée constituée de 16 neurones. Un coefficient de corrélation de Pearson (r) de 0,647, une erreur quadratique moyenne (RMSE) de 0,069 m³/m³, une erreur absolue moyenne (MAE) de 0,05 m³/m³ et une erreur relative moyenne (MRE) de 22,2 % ont été obtenues. Paloscia *et al.* [2013] ont obtenu de meilleurs résultats avec des r de 0,924 et 0,89 et des erreurs quadratiques moyennes de 0,026 et 0,032 avec des RNA utilisant des données provenant d'images acquises en simple polarisation et en double polarisation, respectivement.

La régression linéaire (figure 5.1) permet d'observer la relation entre l'humidité observée (m_{vo}) et estimée (m_{ve}). La pente de 0,39 met en évidence une sous-estimation des données simulées par rapport aux données observées, qui se manifeste particulièrement pour les mesures d'humidité les plus élevées. On note également une dispersion importante, confirmée par les différentes statistiques obtenues, notamment le coefficient de corrélation et le RMSE.

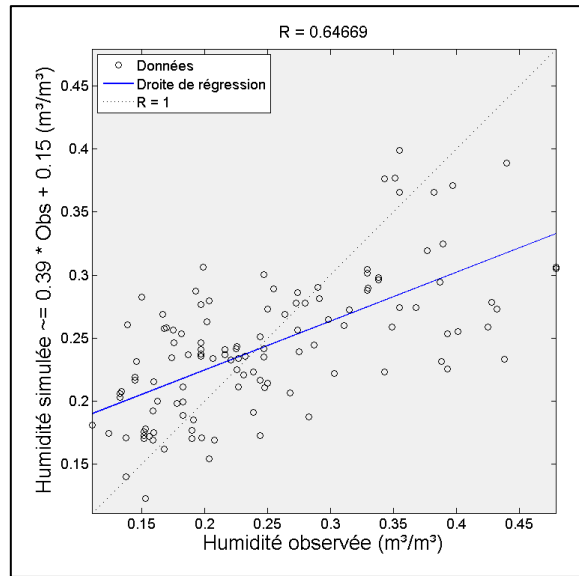


Figure 5.1 Régression linéaire du meilleur RNA avec comme intrants l'angle d'incidence et $\sigma^{\circ}\text{HH}$

En analysant les résidus en fonction de l'humidité du sol observée (figure 5.2) on constate que, de manière générale, le RNA surestime les valeurs inférieures à $0,20 \text{ m}^3/\text{m}^3$ et sous-estime celles supérieures à $0,35 \text{ m}^3/\text{m}^3$. Les résidus les plus importants sont obtenus aux valeurs élevées d'humidité du sol pour lesquelles une importante sous-estimation est observée. Ce résultat s'explique par le fait que le coefficient de rétrodiffusion est proportionnel à l'humidité du sol jusqu'à $0,35 \text{ m}^3/\text{m}^3$, au-delà de laquelle le signal devient insensible à l'humidité du sol [Dobson et Ulaby, 1981]. Lorsque l'humidité du sol excède la saturation, la réflexion spéculaire liée à l'accumulation d'eau en surface augmente et une diminution du signal rétrodiffusé est observée [*Ibid.*].

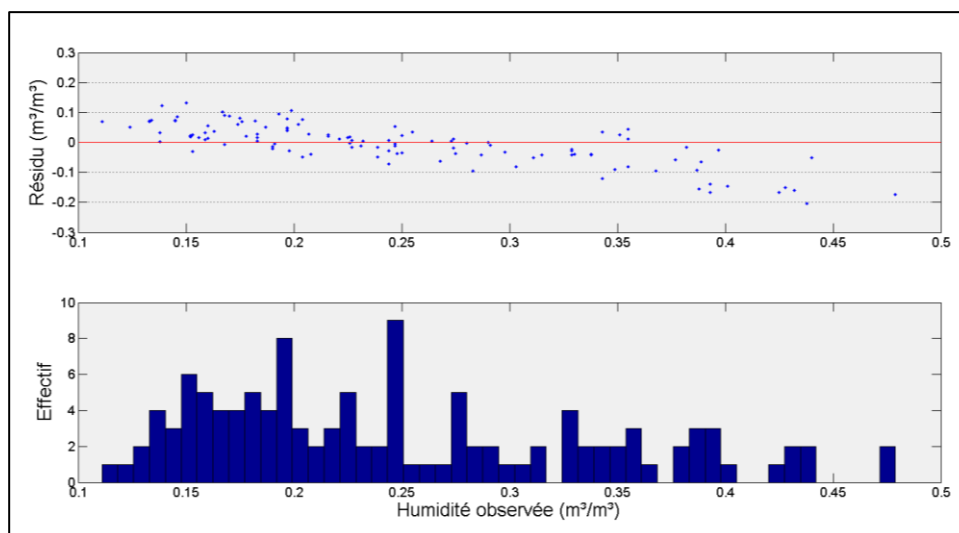


Figure 5.2 Résidus en fonction de l'humidité du sol (haut) et distribution des observations (bas) pour le $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}}$

Cartes d'estimation de l'humidité

Les statistiques des 30 meilleurs RNA obtenus (tableau 5.1) permettent de constater que le coefficient de Pearson moyen et la MRE sont légèrement dégradés comparativement à 0,647 et 22,2% pour le meilleur des 30 RNA. Les valeurs de RMSE et de MAE sont quant à elles très semblables à celles obtenues avec le meilleur RNA. Les faibles valeurs d'écart-type mettent en évidence la faible variabilité entre les différents RNA.

Tableau 5.1 Sommaire des statistiques des 30 meilleurs $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}}$

Variables entrantes	Moyenne				Domaine de valeur							
	Coefficient de Pearson (r)	RMSE (m³/m³)	MAE (m³/m³)	MRE (%)	Coefficient de Pearson (r)		RMSE (m³/m³)		MAE (m³/m³)		MRE (%)	
					Min	Max	Min	Max	Min	Max	Min	Max
θ° et $\sigma^\circ \text{HH}$	0,592	0,073	0,057	24,7	0,55	0,66	0,068	0,076	0,051	0,061	22,0	26,7

Les cartes d'humidité estimée obtenues à angles d'incidence faibles et élevés sont présentées aux figure 5.3 et figure 5.4 respectivement. De manière générale, des résultats différents sont obtenus en fonction de l'angle d'incidence auquel les images RADARSAT-2 ont été acquises.

À angles d'incidence faibles, on observe une sensibilité des RNA à l'humidité du sol. En effet, on remarque un assèchement marqué entre les conditions humides (figure 5.3 a1) et sèches (figure 5.3 b1 et c1). Au début de la saison de croissance, des zones bien définies où

l'humidité estimée est plus élevée sont présentes en conditions sèches (figure 5.3 b1) alors qu'en conditions humides (figure 5.3 a1), ces zones sont moins distinctes.

En superposant la couche des sols ayant un mauvais drainage aux cartes d'humidité moyenne, on constate que les zones les plus humides correspondent généralement à ces sols, mis à part pour les conditions humides du 1^{er} mai (figure 5.3 a2) pour lesquelles le contraste est moins important. Les zones où l'humidité estimée est la plus élevée correspondent donc majoritairement à des milieux humides, principalement des marécages et, dans une moindre mesure, des marais. On note que ces zones interprétées par les RNA comme étant plus humides correspondent précisément à des sols mal drainés (figure 5.4 b2 et c2) principalement recouverts de marécages et de marais, tel que présenté à la figure 4.9. Or, ces zones sont également caractérisées par la présence de végétation arbustive et forestière. Par conséquent, il est possible que les RNA aient décelé la présence d'humidité élevée, alors qu'il pourrait plutôt s'agir d'un effet causé par la présence de la végétation. Il est également possible que les RNA aient effectivement décelé la présence d'humidité du sol élevée par effet de double réflexion du signal radar entre la végétation et l'accumulation d'eau à la surface du sol. Dans ce cas, les RNA permettent de rendre compte d'une humidité du sol élevée sans toutefois être en mesure de la quantifier avec précision.

On observe toutefois une zone distinctement plus sèche avec des limites rectilignes dans les marais de Holland sur la carte du 25 mai 2012. Les limites sud et est (mises en évidence à la figure 5.3 b2) correspondant au canal de drainage, on peut supposer que cette zone correspond à des champs drainés, considérant que les valeurs de σ_{HH}° y sont plus faibles. Ce phénomène n'est toutefois pas observé sur la carte du 31 juillet 2013. Les précipitations plus importantes enregistrées durant la saison 2013 ont contribué à maintenir une humidité du sol plus élevée qu'en 2012, ce qui pourrait avoir diminué l'efficacité du réseau de drainage.

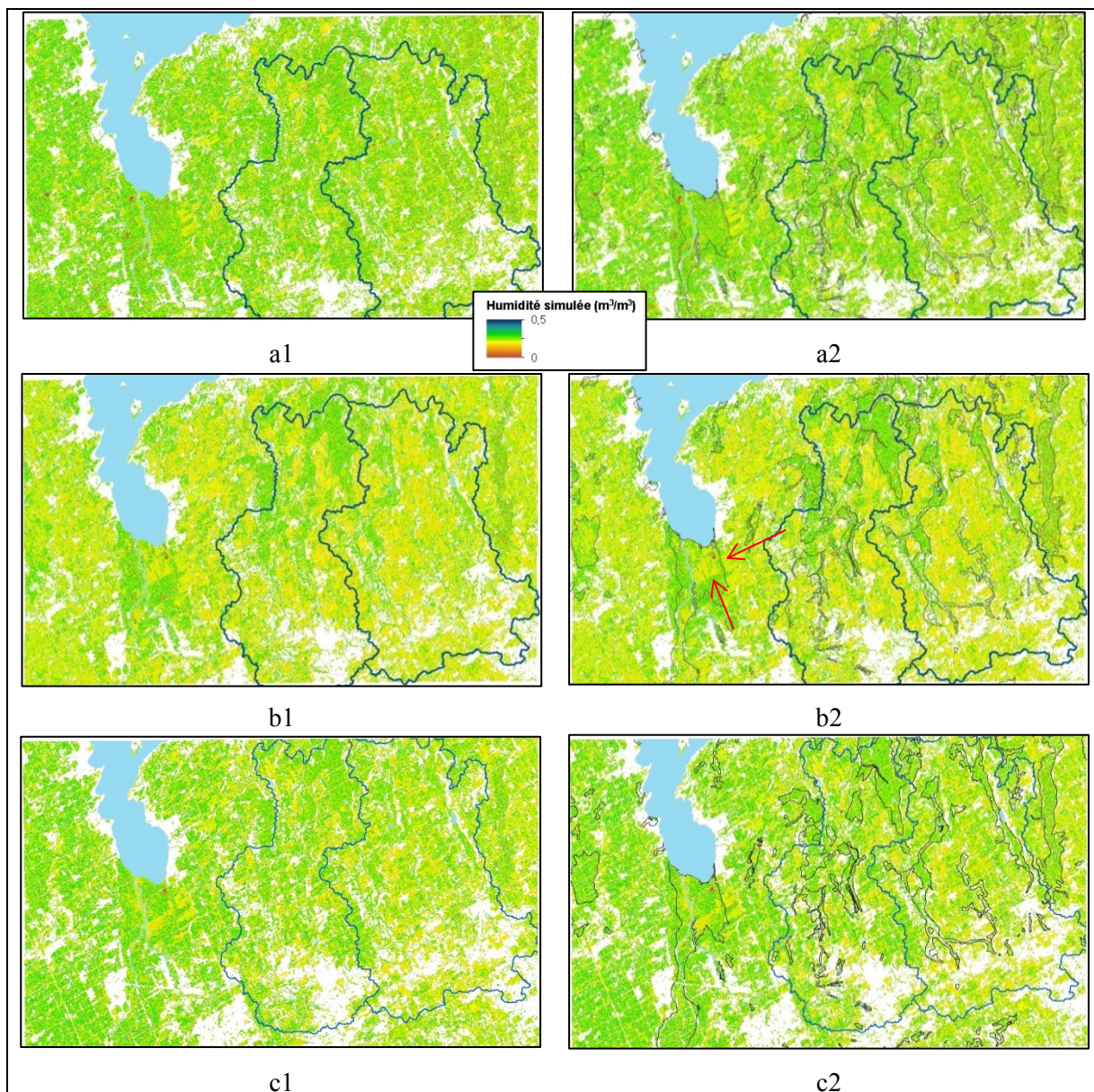


Figure 5.3 Carte d'estimation de l'humidité du sol avec $RNA_{\theta^{\circ}, \sigma^{\circ}HH}$ à faibles angles d'incidence et superposition des sols très mal drainés pour le 1^{er} mai 2012 – conditions humides (a1 et a2), le 25 mai 2012 – conditions sèches (b1 et b2) et le 31 juillet 2013 – conditions sèches (c1 et c2)

À angles d'incidence élevés, l'humidité estimée en conditions sèches (figure 5.4 a1) est légèrement plus élevée qu'en conditions humides (figure 5.4 b1) malgré des valeurs d'humidité du sol observées (m_{vo}) de 0,134 et 0,244 m³/m³ pour le 15 juillet et le 8 août, respectivement. La végétation abondante et les angles d'incidence élevés sont probablement responsables de ces résultats contradictoires, comme il le sera expliqué plus loin.

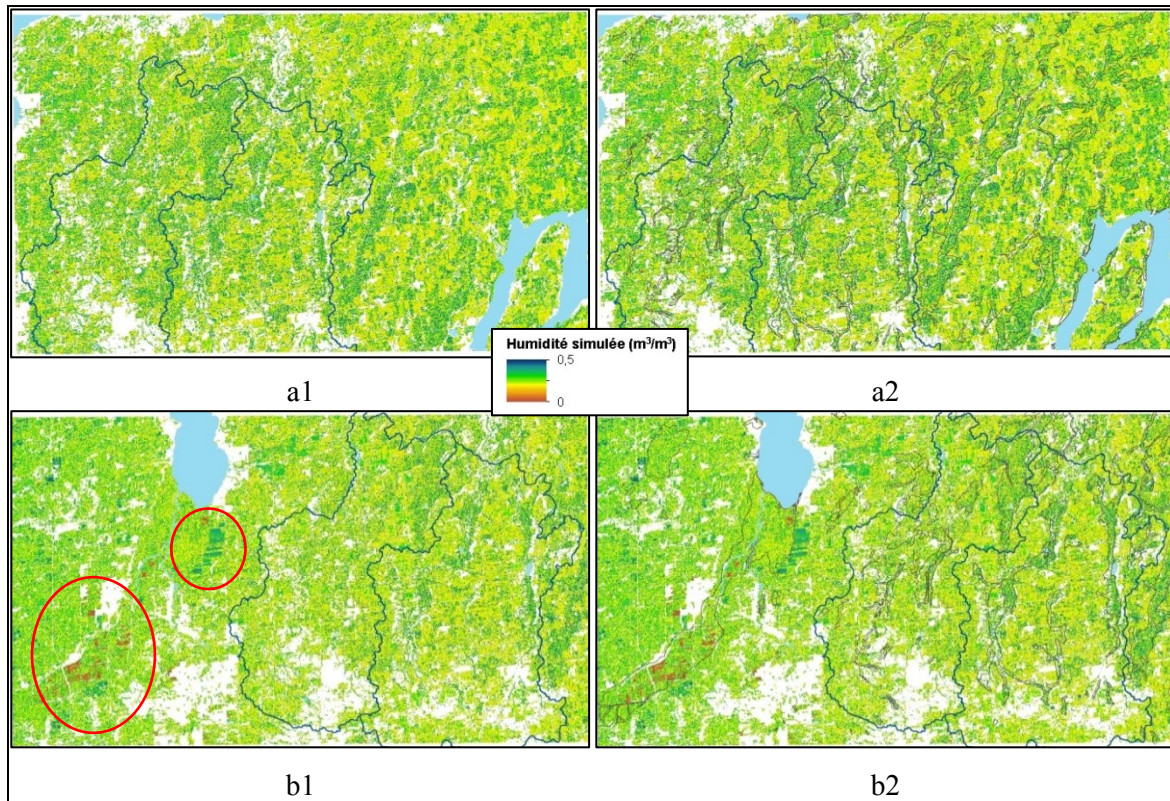


Figure 5.4 Carte d'estimation de l'humidité du sol avec $\text{RNA}_{\theta^\circ}$, σ_{HH}° et superposition des sols très mal drainés à angles d'incidence élevés pour le 15 juillet 2012 – conditions sèches (a1 et a2) et le 8 août 2012 – conditions humides (b1 et b2)

De plus, les RNA ne sont pas en mesure d'interpréter correctement les valeurs de σ_{HH}° qui se situent en dehors du domaine de valeurs du jeu de données d'entraînement. Ces zones sont mises en évidence sur la figure 5.4 b1. Dans ces situations, le RNA se trouve donc à faire de l'extrapolation, tâche pour laquelle les RNA ne sont pas toujours performants. L'incertitude associée à ces zones est d'ailleurs très élevée.

En analysant les sous-bassins Baldwin et Pepperlaw, les histogrammes de la figure 5.5 ont été obtenus. Pour des angles d'incidence faibles et en présence de végétation peu développée (1^{er} et 25 mai 2012), on observe des différences marquées entre les histogrammes en conditions sèches et humides. La probabilité d'observer une humidité estimée entre 0,10 et 0,20, notée $P(0,10 < x < 0,20)$, est 16,7 % plus élevée alors que $P(0,20 < x < 0,30)$ et $P(x > 0,30)$ sont inférieures de 12,5 % et 3,6 %, respectivement, en conditions sèches comparativement aux conditions humides. Ces résultats s'expliquent par une sensibilité relativement élevée du signal radar à l'humidité du sol compte tenu de l'angle d'incidence de 18-21° et la faible

végétation présente. En effet, Ulaby et Batlivala [1976] ont identifié qu'un angle d'incidence entre 7-15° était optimal pour minimiser l'effet de la rugosité et maximiser la sensibilité à m_v pour les sols nus.

Lorsque la végétation est bien développée et que l'angle d'incidence demeure faible (31 juillet 2013), l'ensemble des probabilités se situent à environ la moitié de l'écart entre conditions sèches et humides sans végétation, malgré une humidité observée identique à celle en conditions sèches. Cet écart peut s'expliquer par le fait que 51 mm de pluie ont été enregistrés à la station météorologique la plus proche pour la journée du 31 juillet 2013. Ainsi, malgré les faibles angles d'incidence, le développement avancé des cultures et la présence probable d'eau sur la végétation a généré une rétrodiffusion plus importante pour les zones cultivées, ce qui s'est traduit par une humidité plus élevée. Ainsi, l'influence de la végétation sur la rétrodiffusion est réduite lorsque les angles d'incidence sont faibles. L'humidité du sol estimée dans ces conditions est toutefois probablement davantage liée à la végétation qu'au sol.

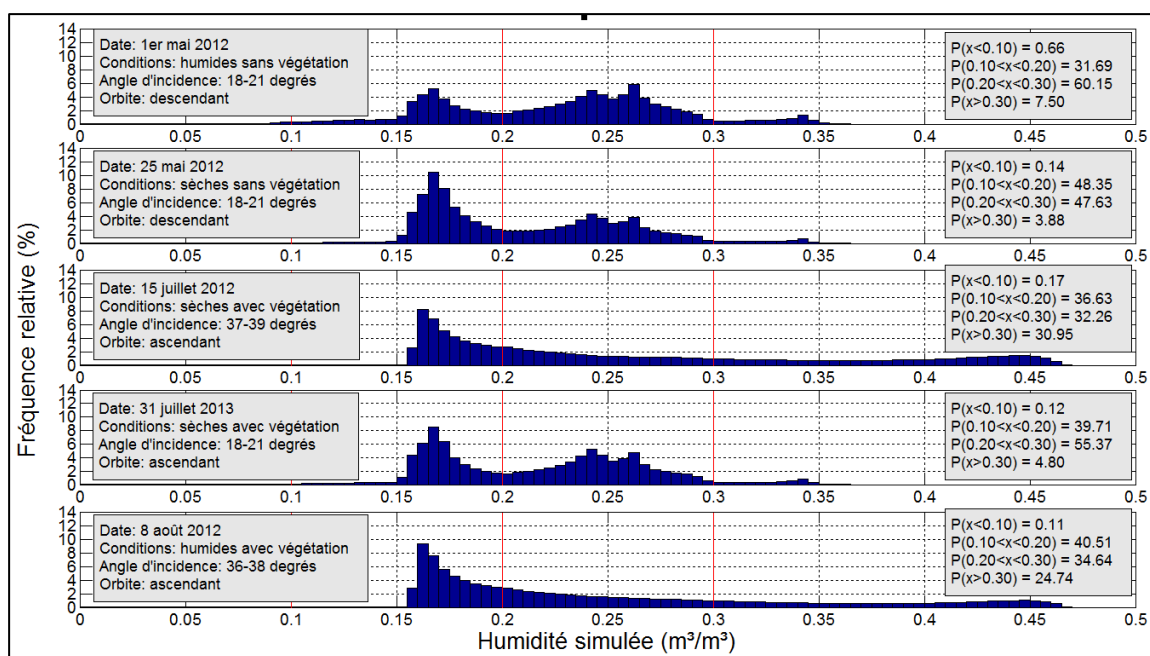


Figure 5.5 Histogrammes de l'humidité estimée avec $RNA_{\theta^{\circ}, \sigma^{\circ}HH}$ pour les sous-bassins versants Baldwin et Pepperlaw

Lorsque la végétation est abondante et les angles d'incidence sont élevés (cartes du 15 juillet 2012 et du 8 août 2012), les histogrammes en conditions sèches et humides sont très semblables. De plus, la dispersion des valeurs d'humidité est contraire à ce qui était attendu.

En effet, en conditions sèches $P(0,10 < x < 0,20)$ est plus faible et $P(x > 0,30)$ plus élevée qu'en conditions humides. Aussi, les probabilités $P(x > 0,30)$ sont très élevées considérant que les observations sont de 0,134 et 0,244 m³/m³ pour le 15 juillet et le 8 août, respectivement.

La végétation bien développée ainsi que les angles d'incidence élevés lors de l'acquisition des images du 15 juillet et 8 août 2012 expliquent les résultats obtenus. En effet, la végétation influence de manière importante le signal rétrodiffusé de par son effet sur la diffusion de volume qu'elle entraîne. À mesure que celle-ci croît, sa contribution au coefficient de rétrodiffusion total augmente jusqu'à devenir la composante dominante. La zone couverte par ces images est un territoire où les principales occupations du sol sont les terres agricoles, principalement du maïs, ainsi que les cultures pérennes et les pâturages. Il a été démontré que dans les cultures de maïs de plus de 50 cm, taille atteinte vers la mi-juin, la contribution de la diffusion de surface est inférieure à 50 %. De plus, à des angles d'incidence élevés (34° à 40°) les mêmes résultats ont été obtenus pour des cultures de maïs de moins de 50 cm [Trudel *et al.*, 2012]. De nombreuses études ont démontré que la contribution de la rugosité de surface et la quantité de biomasse perçue augmentent avec une augmentation de l'angle d'incidence [Baghdadi *et al.*, 2002b; Paloscia, 2002; Romshoo *et al.*, 2002], rendant ainsi l'estimation de l'humidité du sol pratiquement impossible. Considérant que les images du 15 juillet et du 8 août 2012 ont été acquises à des angles d'incidence variant entre 36 et 39°, il est raisonnable de croire que l'information contenue dans σ°_{HH} provient essentiellement de la végétation. La différence marquée entre l'histogramme obtenu à faible angle d'incidence (31 juillet 2013) et ceux à angles élevés (15 juillet et 8 août 2012) ainsi que la similarité entre ceux obtenus à angles faibles, indépendamment de la quantité de végétation présente, tendent à valider cette hypothèse. Notons que malgré que les images aient été acquises sur deux années, il est raisonnable de croire que les possibles changements de culture ne peuvent à eux seuls expliquer la différence entre les histogrammes obtenus à des angles d'incidence élevés et ceux à angles faibles.

Cartes d'incertitude sur l'estimation

En observant la répartition spatiale de l'incertitude (figure 5.6), on constate qu'elle n'est pas corrélée à l'humidité estimée et qu'elle est répartie de manière plus homogène que cette dernière. En effet, les zones très mal drainées où l'humidité estimée est la plus élevée ne correspondent pas nécessairement à des zones d'incertitude très élevée.

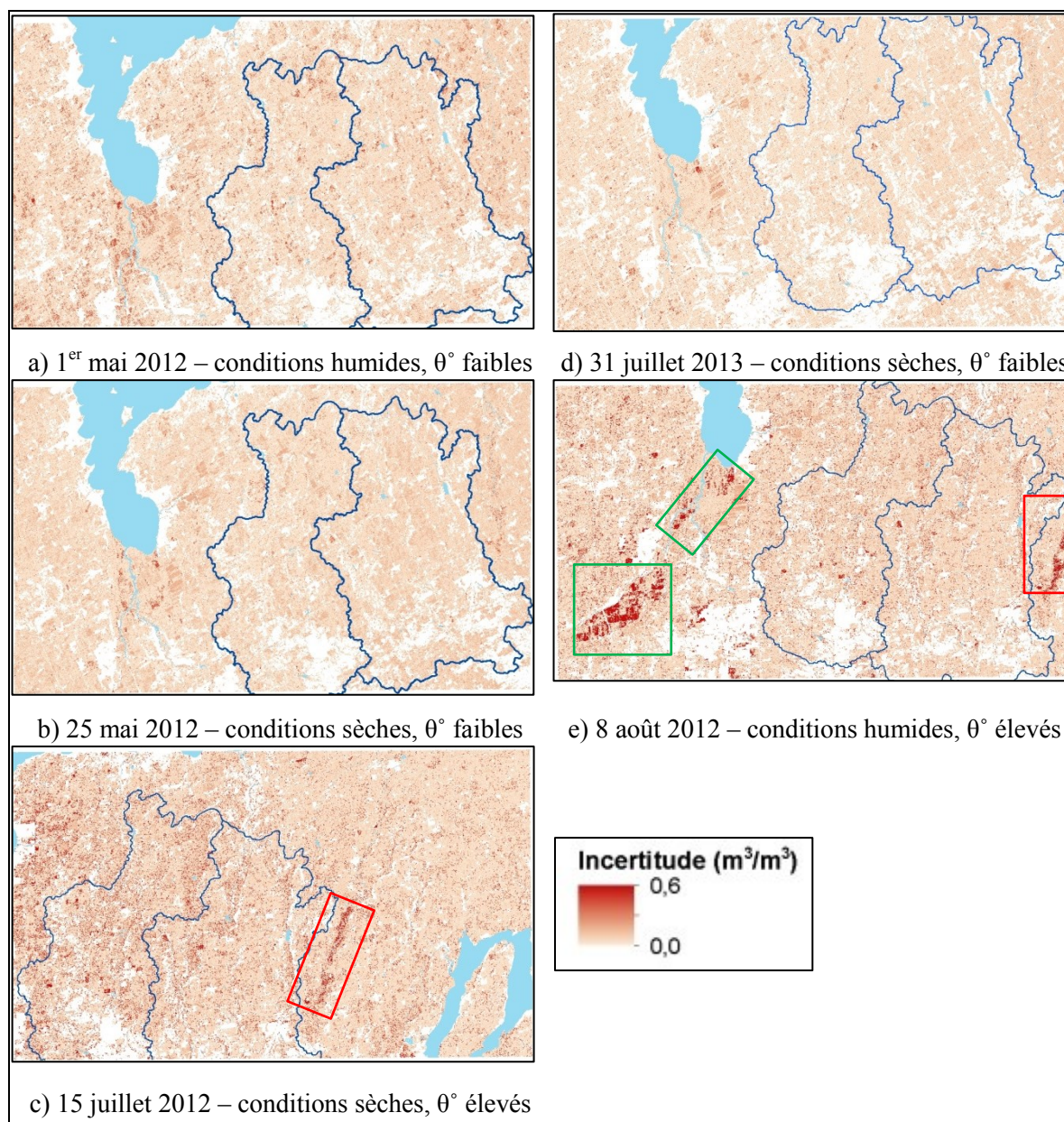


Figure 5.6 Cartes d'incertitude sur l'estimation de l'humidité du sol obtenues avec $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}}$

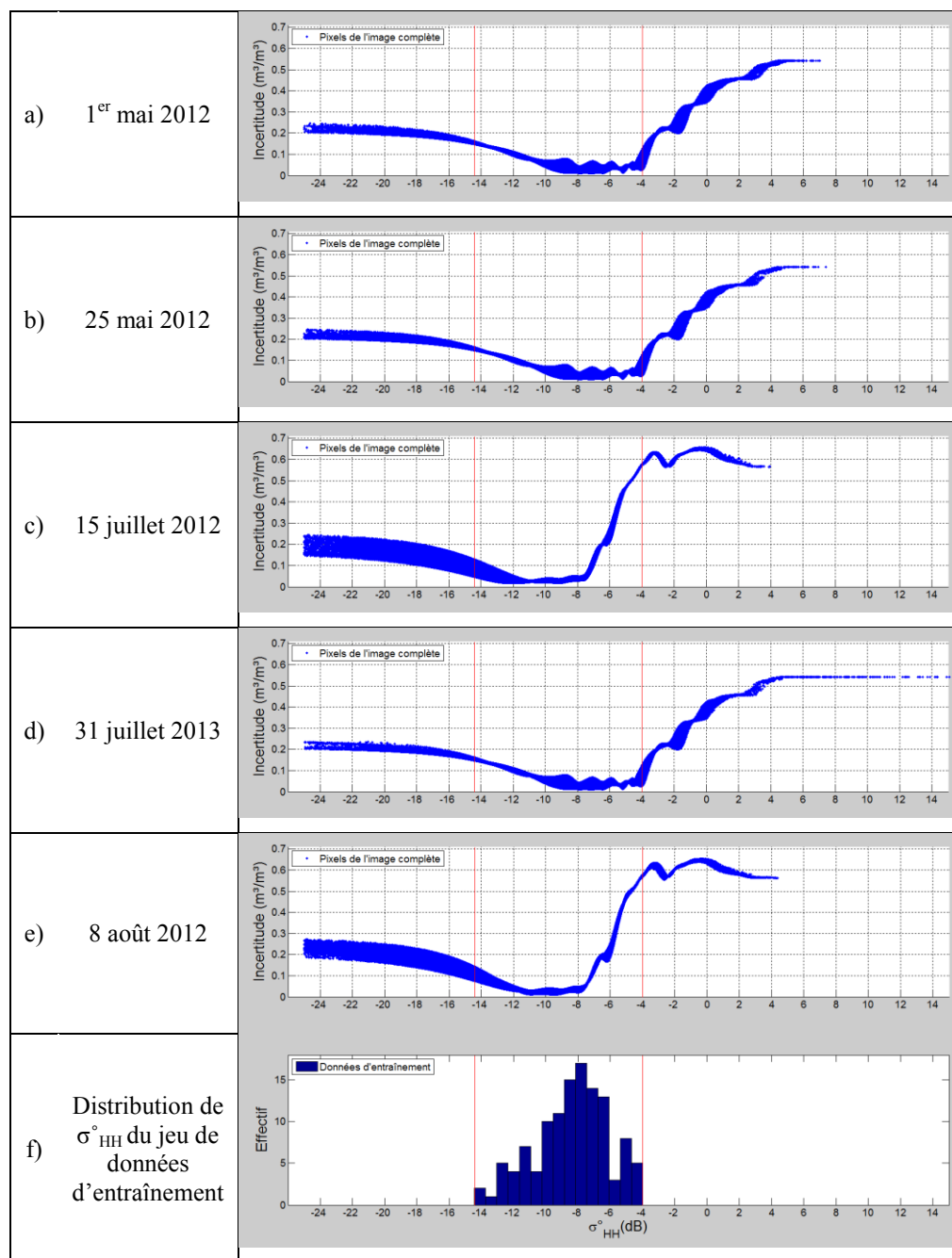
Comme ces zones sont constituées principalement de marais et de marécages et qu'une végétation ligneuse y est présente de manière quasi-permanente, le coefficient de rétrodiffusion est relativement constant dans le temps. Toutefois, lorsque la nappe phréatique se trouve près de la surface, la double réflexion augmente et génère des coefficients de rétrodiffusion nettement plus élevés. Ces valeurs n'étant pas présentes dans le jeu de données d'entraînement, une variabilité élevée de l'humidité du sol estimée par les 30 RNA résulte en une incertitude élevée dans ces conditions. Ainsi, l'incertitude obtenue pour les zones mal

drainées est davantage fonction de la végétation et de la hauteur de la nappe phréatique. L'ajout de σ_{VV}° , moins sensible à ce phénomène, comme intrant aux RNA aurait pu contribuer à réduire les ambiguïtés dans ces zones. Par contre, aucune donnée représentative de ces conditions n'aurait été présente dans le jeu de données utilisé pour le développement des RNA puisqu'aucune sonde d'humidité n'a été installée dans ces zones.

Pour les angles d'incidence faibles (figure 5.7 a, b et d), l'incertitude est généralement plus faible alors que pour les angles élevés (figure 5.7 c et e), il y a présence de zones relativement bien définies où l'incertitude est nettement plus élevée. La figure 5.7 met en évidence que la magnitude de l'incertitude est plutôt fonction du coefficient de rétrodiffusion et de l'angle d'incidence.

En effet, les nuages de points représentant l'incertitude en fonction de σ_{HH}° sont de forme parabolique pour toutes les images d'humidité du sol estimée. Ce résultat démontre l'importance que le jeu de données d'entraînement puisse couvrir l'ensemble des conditions sur lesquelles le RNA sera appliqué. En effet, lorsque les valeurs de σ_{HH}° fournies aux RNA sortent du domaine de valeurs utilisées lors de l'entraînement (figure 5.7 f), une augmentation marquée de l'incertitude est observée. D'autre part, l'influence du nombre de données sur l'incertitude est aussi évidente : c'est autour de - 8 dB que l'incertitude est la plus faible, ce qui correspond aux valeurs ayant l'effectif le plus important. La valeur supérieure de σ_{HH}° à partir de laquelle l'incertitude augmente drastiquement est elle aussi fonction du jeu d'entraînement. En effet, à angles d'incidence faibles, c'est au-delà de -4 dB et -8 dB à forts angles d'incidence. Ces valeurs correspondent aux valeurs maximales de σ_{HH}° associées à des angles d'incidence de plus ou moins 30° présentes dans le jeu de données. L'incertitude atteint un plateau à 0,6 m³/m³ à 4 dB et -4 dB pour les angles d'incidence faibles et élevés, respectivement. Les RNA produisent donc des valeurs de m_{ve} irréalistes très faibles ou très élevées lorsque le signal rétrodiffusé est fort et supérieur aux valeurs maximales de σ_{HH}° présentées lors de l'entraînement.

Dans le cas des faibles valeurs de σ_{HH}° , l'incertitude augmente lorsque σ_{HH}° est inférieur à -10 dB et -12 dB pour les angles d'incidence faibles et élevés, respectivement et atteint un plateau à 0,2 m³/m³ à -14 dB. L'incertitude est donc beaucoup moins sensible aux faibles valeurs de σ_{HH}° .

Figure 5.7 Incertitude en fonction de σ°_{HH}

L'incertitude particulièrement élevée observée sur les cartes du 15 juillet et du 8 août 2012 est liée à un marais et à des champs de cultures maraîchères (encadrés rouge et verts, respectivement, à la figure 5.6 c et e). Dans ces zones, le coefficient de rétrodiffusion est supérieur à -6 dB. Les facteurs suivants ont contribué à ces valeurs élevées : 1) les importantes quantités de pluie précédant l'acquisition des images RSO; 2) les forts angles d'incidence de ces images. En particulier, plus de 52 mm de pluie ont été enregistrés à la station

météorologique la plus proche au cours de la semaine précédant l'acquisition du 8 août, dont 30 mm le 4 août. Ces précipitations ont probablement fait augmenter le niveau d'eau du marais et produit des accumulations d'eau dans les champs, ce qui a contribué à une importante double réflexion. L'influence de l'angle d'incidence est également mise en évidence, puisque malgré les 51 mm de pluie enregistrés le 31 juillet 2013 - dont une fraction est vraisemblablement tombée avant l'acquisition de l'image RADARSAT à faible angle d'incidence - ces zones d'incertitude ne sont pas présentes. Ces valeurs de σ_{HH}° n'étant pas présentes dans le jeu de données utilisé lors de la phase d'entraînement, les RNA n'ont pas appris ce cas de figure. Par conséquent, la valeur générée par chacun des 30 RNA est très variable, ce qui a généré une importante incertitude.

L'analyse de la dispersion de l'incertitude dans les sous-bassins versants Baldwin et Pefferlaw (figure 5.8) confirme les observations faites sur l'ensemble des images. En effet, on remarque que la dispersion est principalement fonction de l'angle d'incidence.

D'une part, les histogrammes du 1^{er} mai 2012, du 25 mai 2012 et du 31 juillet 2013 sont très semblables et présentent une incertitude minimale et maximale plus faibles que celles du 15 juillet 2012 et du 8 août 2012. D'autre part, l'incertitude à angles d'incidence forts est très concentrée autour de $0,025 \text{ m}^3/\text{m}^3$ avec environ 40 % des pixels, contrairement à une distribution plus uniforme entre $0,01$ et $0,05 \text{ m}^3/\text{m}^3$ à angles faibles.

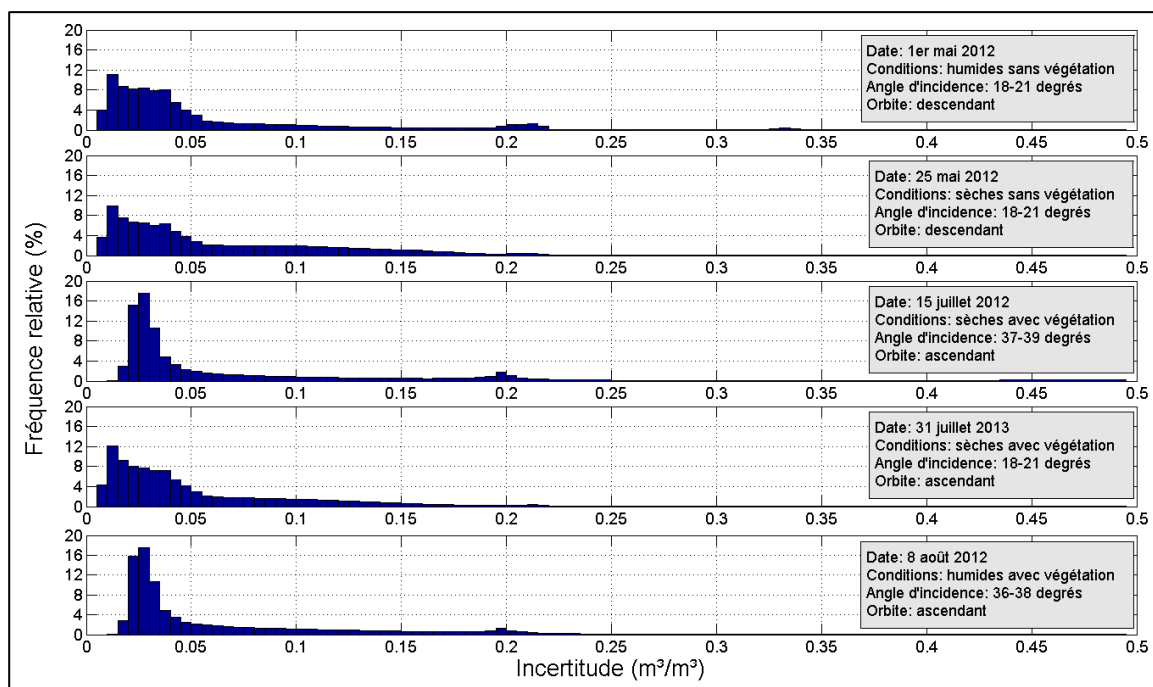


Figure 5.8 Histogrammes d'incertitude pour les sous-bassins versants Baldwin et Pefferlaw pour les $RNA_{\theta^{\circ}, \sigma^{\circ}HH}$

On peut donc conclure que les RNA utilisant l'angle d'incidence (θ°) et le coefficient de rétrodiffusion radar en polarisation HH ($\sigma^{\circ}HH$) comme intrants sont sensibles aux variations d'humidité du sol et sont donc en mesure de l'estimer. Cette capacité est toutefois limitée aux images acquises à des angles d'incidence faibles. Dans ces conditions, l'incertitude sur les estimations est majoritairement inférieure à $0,05 \text{ m}^3/\text{m}^3$, et ce, indépendamment des conditions d'humidité du sol. Lorsqu'il y a présence de végétation abondante, la fiabilité de l'humidité estimée est variable en fonction des caractéristiques de la végétation. En effet, comme la densité et la géométrie de la végétation influencent la capacité de pénétration du signal radar et qu'aucune mesure in situ n'a été réalisée, il est difficile d'évaluer la contribution de l'humidité du sol à $\sigma^{\circ}HH$ et, par le fait même, la fiabilité de l'humidité estimée. On peut penser que dans les champs de cultures maraîchères, le signal radar contient davantage d'information provenant de la surface du sol. En effet, le feuillage peu important des oignons et des carottes (70 % de la superficie cultivée en 2006 [Planscape, 2009]) et la faible taille des plants permet une pénétration plus importante du signal radar, comparativement à des cultures comme le maïs. Ainsi, compte tenu de la faible quantité de végétation présente lors de l'acquisition des images du 1^{er} et 25 mai 2012, l'estimation de l'humidité qui a été faite à partir de ces données

peut donc être jugée fiable. Dans le cas du 31 juillet 2013, la fiabilité de l'humidité estimée est variable en fonction des caractéristiques de la végétation.

Toutefois, la hauteur de la nappe phréatique dans les milieux humides est un facteur limitant puisque lorsque celle-ci est trop haute, la réflexion spéculaire ainsi que la double réflexion seront trop importantes pour obtenir de l'information provenant directement du sol. Par contre, l'humidité du sol élevée produite par les RNA s'approche de la réalité puisque, dans ces conditions, le sol est saturé d'eau. La présence de végétation est donc un avantage puisque sans elle, la réflexion spéculaire domine et l'humidité du sol risque d'être sous-estimée de manière importante. Toutefois, lorsque la nappe phréatique est trop basse pour produire de la double réflexion, la rétrodiffusion captée ne contient que peu d'information provenant de la surface. Compte tenu que des milieux humides sont présents parmi les sols très mal drainés (voir section 0), une proportion inconnue de l'humidité estimée dans ces zones n'est pas fiable, principalement en conditions sèches.

5.1.2 Angle d'incidence, coefficient de rétrodiffusion HH et pente

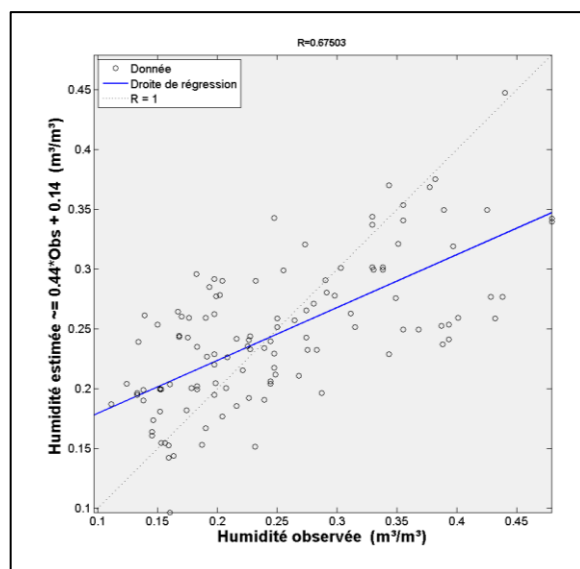
Dans le but de prendre en compte la topographie et son influence sur l'humidité du sol, la pente a été ajoutée comme troisième variable entrante lors de l'entraînement des RNA.

Le meilleur RNA obtenu avec cette combinaison d'intrants possède une couche cachée constituée de 13 neurones. Au niveau statistique, l'ajout de la pente a résulté en une diminution non significative de l'ensemble des critères de performance (tableau 5.2).

Tableau 5.2 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double

Variables entrantes	Coefficient de Pearson (r)	RMSE (m ³ /m ³)	MAE (m ³ /m ³)	MRE (%)
θ° et σ°_{HH}	0,647	0,069	0,053	22,2
θ° , σ°_{HH} et pente	0,639	0,071	0,054	25,5

La régression linéaire (figure 5.9) permet d'observer la relation entre l'humidité observée et celle simulée. La pente de 0,44 met en évidence une sous-estimation des données simulées par rapport aux données observées qui se manifeste particulièrement pour les mesures d'humidité les plus élevées.

Figure 5.9 Régression linéaire du $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}HH, \text{pente}}$

Les erreurs commises sur l'estimation peuvent être divisées en deux classes, selon si l'observation est inférieure ou supérieure à $0,27 \text{ m}^3/\text{m}^3$. Pour les valeurs inférieures à $0,27 \text{ m}^3/\text{m}^3$, l'humidité simulée est davantage surestimée et la magnitude des erreurs plus grande que lorsque sous-estimée. Au-delà de $0,27 \text{ m}^3/\text{m}^3$, le RNA sous-estime de manière importante l'humidité du sol dans la quasi-totalité des cas. En considérant les valeurs des critères de performance obtenues ainsi que la dispersion des erreurs, les performances de ce RNA sont très semblables à celui ayant l'angle d'incidence et σ°_{HH} comme intrants.

Cartes d'estimation de l'humidité du sol et de l'incertitude

L'ajout de la pente comme variable entrante à produit d'importants changements sur les cartes d'humidité estimée (figure 5.10). Lorsque comparées à celles obtenues avec uniquement σ°_{HH} et l'angle d'incidence, on observe une diminution de m_{ve} dans les zones les plus sèches et une hausse de m_{ve} dans les zones les plus humides. Le contraste est le plus important en conditions de végétation abondante pour les angles d'incidence élevés (figure 5.10 c et e). L'humidité estimée dans les endroits ayant une pente très forte est très faible, ce qui traduit bien l'important ruissellement de surface qui s'y produit. Par contre, l'utilisation de la pente produit des artefacts dans les cartes d'humidité et d'incertitude. En effet, en observant à une plus petite échelle, tel que présenté à la figure 5.11, on distingue clairement les isolignes produites lors de la conversion du MNÉ en données de pente. De plus, l'incertitude associée aux zones de forte pente est élevée.

Finalement, en analysant l'incertitude pour les sous-bassins versant Baldwin et Pepperlaw (figure 5.12), on observe un pourcentage de valeurs élevées nettement plus important comparativement aux $RNA_{\theta^\circ, \sigma^\circ HH}$. En effet, l'incertitude pour les $RNA_{\theta^\circ, \sigma^\circ HH, pente}$ est majoritairement inférieure à $0,1 \text{ m}^3/\text{m}^3$ et $0,2 \text{ m}^3/\text{m}^3$ à angles d'incidence faibles et élevés, respectivement, comparativement à une majorité inférieure à $0,05 \text{ m}^3/\text{m}^3$ indépendamment de l'angle d'incidence pour les $RNA_{\theta^\circ, \sigma^\circ HH}$ (voir la figure 5.8).

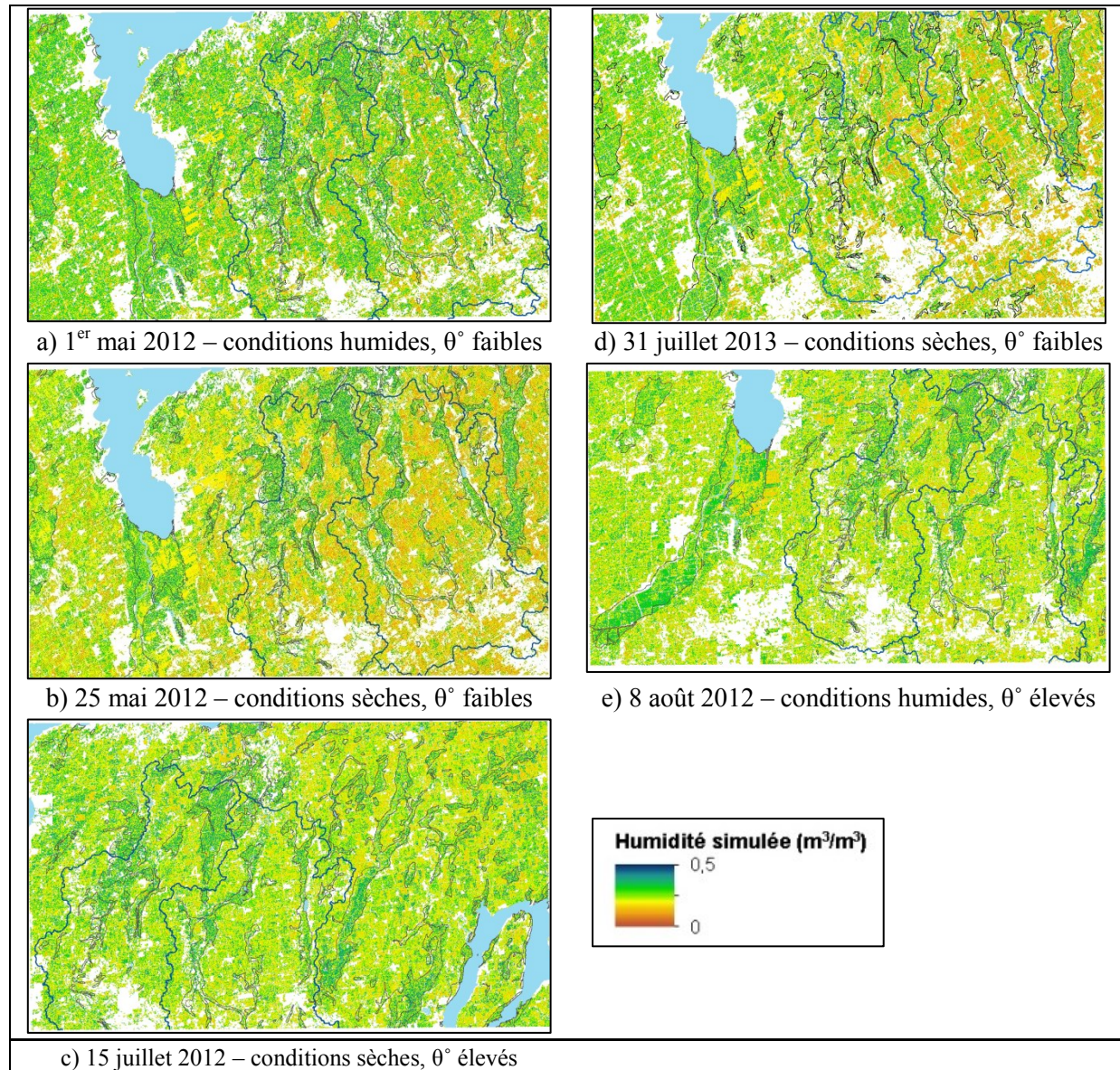


Figure 5.10 Carte d'estimation de l'humidité du sol avec avec $RNA_{\theta^\circ, \sigma^\circ HH, pente}$

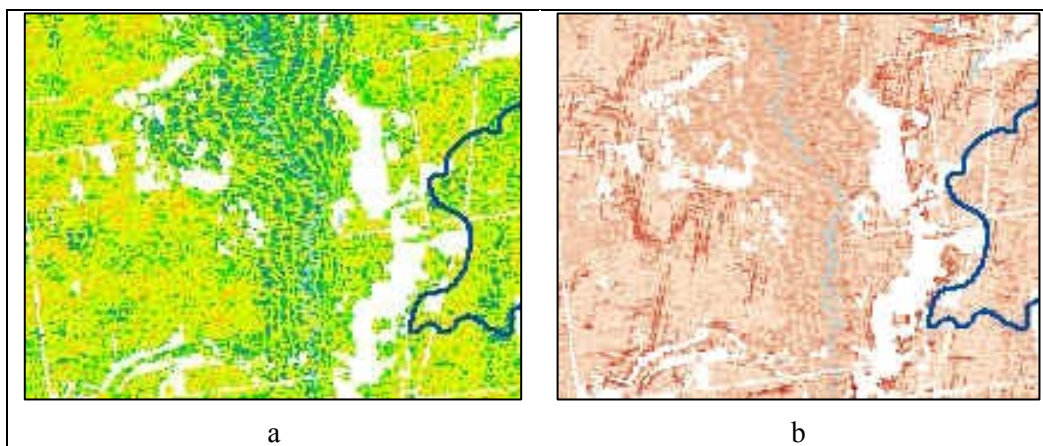


Figure 5.11 Agrandissement d'une zone de la carte d'humidité estimée (a) et d'incertitude (b) obtenues pour le 15 juillet 2012

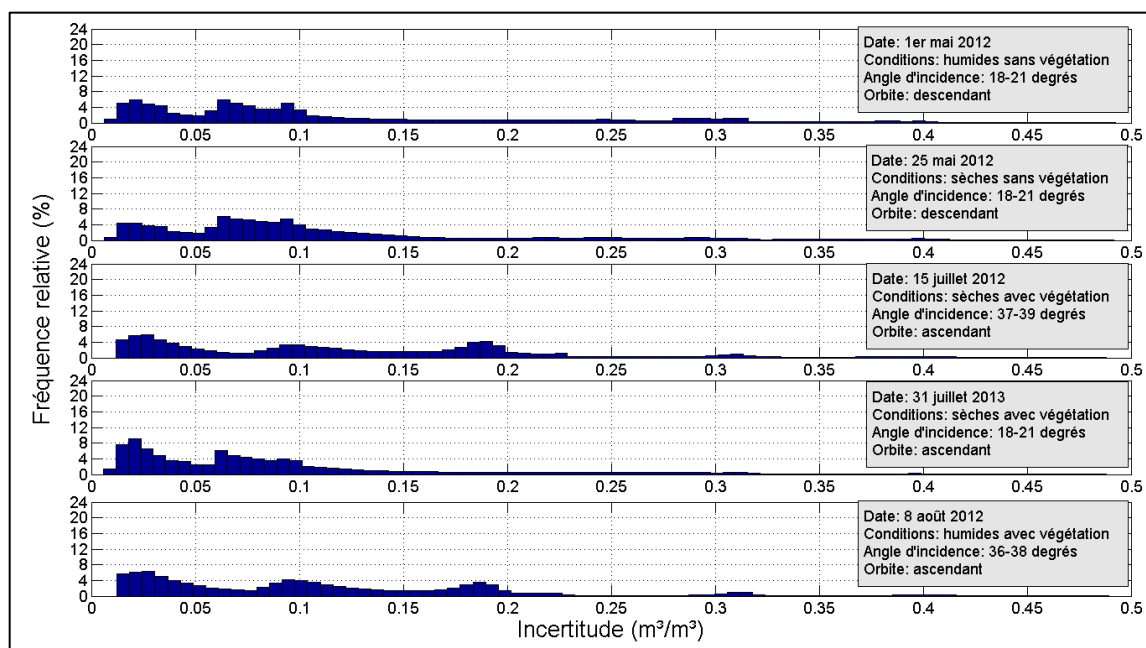


Figure 5.12 Histogrammes d'incertitude pour les sous-bassins versants Baldwin et Pepperlaw avec $RNA_{\theta, \sigma^* HH, pente}$

Ce résultat s'explique partiellement par le fait que les valeurs des données de pente présentes dans le jeu de données d'entraînement varient entre 1 et 7 % alors que pour l'ensemble de la zone couverte par les cartes, les valeurs de pente peuvent atteindre 122 %. Par contre, tel que présenté à l'annexe b, la pente est supérieure à 10 % sur seulement 15,4 % de la portion terrestre des bassins versant Baldwin et Pepperlaw. Les RNA extrapolent donc sur une faible proportion du territoire.

Compte tenu des performances similaires du meilleur RNA à celui utilisant l'angle d'incidence et le coefficient de rétrodiffusion HH, des artefacts présents, mais surtout à cause de l'incertitude beaucoup plus élevée, la pente n'est pas considérée comme une variable permettant d'améliorer la capacité des RNA à estimer l'humidité du sol.

5.1.3 Angle d'incidence, coefficient de rétrodiffusion HH et texture

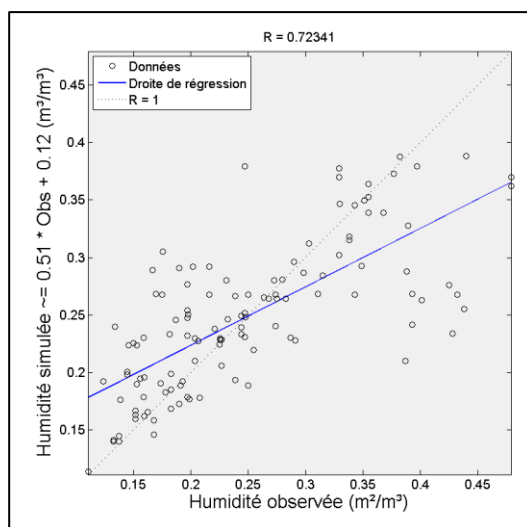
Dans le but de prendre en compte l'effet de la texture du sol sur la perméabilité du sol et la constante diélectrique, les fractions de sable et d'argile ont été ajoutées comme variables entrantes aux réseaux.

Le réseau ayant permis d'obtenir les meilleures performances possède une couche cachée composée de 15 neurones et, en se basant sur les critères de performance, est le plus performant des trois meilleurs RNA obtenus jusqu'à présent (tableau 5.3). Le coefficient de Pearson et l'erreur relative moyenne sont les critères pour lesquels les différences sont les plus marquées. L'erreur relative moyenne inférieure de 4,4 % par rapport à $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}}$ met en évidence une plus faible dispersion des résidus. Toutefois, les différences observées ne sont pas significatives du point de vue statistique ($p\text{-value} > 0,05$).

Tableau 5.3 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double

Variables entrantes	Coefficient de Pearson (r)	RMSE (m ³ /m ³)	MAE (m ³ /m ³)	MRE (%)
θ° et $\sigma^\circ \text{HH}$	0,647	0,069	0,053	22,2
θ° , $\sigma^\circ \text{HH}$ et pente	0,639	0,071	0,054	25,5
θ° , $\sigma^\circ \text{HH}$ et texture	0,723	0,062	0,043	17,8

Bien que la pente soit plus élevée, 0,51 comparativement à 0,39 avec uniquement l'angle d'incidence et $\sigma^\circ \text{HH}$ comme intrants, la régression linéaire (figure 5.13) présente les mêmes tendances. En effet, les faibles valeurs observées sont généralement surestimées et, à l'inverse, les mesures d'humidité élevées sont sous-estimées et plus fortement au-delà de 0,35 m³/m³.

Figure 5.13 Régression linéaire du meilleur $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}HH, \text{texture}}$

Cartes d'estimation de l'humidité et d'incertitude

Les valeurs des critères de performance portent à croire que l'utilisation de la texture comme intrant permet d'améliorer de manière non négligeable la précision de l'estimation de l'humidité du sol. L'analyse des cartes de texture (figure 5.14) conjointement à celles d'estimation de l'humidité du sol obtenues (figure 5.15) ne permet toutefois pas d'arriver à cette conclusion. Bien que l'humidité du sol estimée soit généralement plus élevée pour les sols très mal drainés, la répartition spatiale de l'humidité estimée est fortement corrélée à celle de la texture. L'influence de la fraction de sable sur l'humidité est toutefois largement dominante, excepté pour les valeurs de fraction d'argile très élevées, principalement pour le marais de Holland (voir section 0), où elle atteint 60 %.

Pour les angles d'incidence faibles (figure 5.15 a, b et d), l'humidité estimée augmente généralement avec une diminution de la fraction de sable et une augmentation de la fraction d'argile, ce qui était attendu compte tenu que le sable favorise la percolation alors que l'argile la limite. Par contre, l'humidité estimée augmente avec une augmentation des fractions de sable et d'argile lorsque les angles d'incidence sont élevés (figure 5.15 c et e).

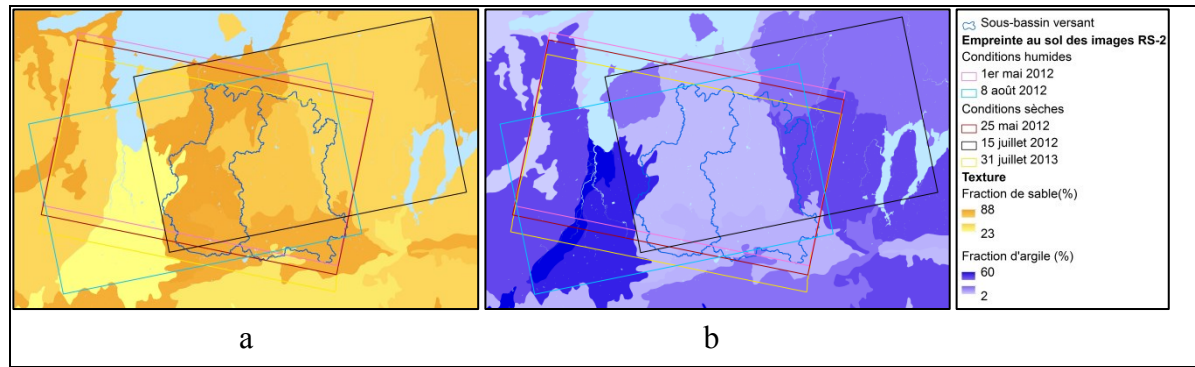


Figure 5.14 Cartes des fractions de sable (a) et d'argile (b)

Plusieurs facteurs peuvent avoir contribué à ces résultats. Premièrement, il est possible que les poids attribués aux fractions de sable et d'argile par les RNA soient trop importants par rapport à celui du coefficient de rétrodiffusion. Il est possible qu'utiliser uniquement la fraction de sable ou d'argile diminuerait la redondance dans les données entrantes et permettrait d'obtenir de meilleurs résultats. Deuxièmement, les données de texture utilisées proviennent de cartes pédologiques recompilées à une échelle de 1:1 000 000. Il est en effet peu probable que d'aussi grandes zones possèdent une texture homogène. Finalement, comme pour une même humidité du sol, σ° est variable en fonction de la texture, il semble que les RNA n'ont pas été en mesure de capter ce lien lors du processus d'entraînement. La prise de mesures de texture à l'emplacement des sondes afin d'avoir des valeurs réelles dans le jeu de données d'apprentissage aurait pu permettre un meilleur apprentissage.

L'incertitude obtenue est nettement plus élevée pour toutes les conditions d'acquisition et d'humidité du sol. En effet, bien que l'incertitude maximale soit moins élevée comparativement aux $RNA_{\theta^\circ, \sigma^\circ HH}$, l'incertitude médiane est deux fois plus élevée dans les sous-bassins Baldwin et Pepperlaw. L'incertitude est principalement corrélée à la fraction d'argile, ce qui s'explique par le domaine de valeurs très restreint du jeu de données d'entraînement. En effet, le domaine de valeurs de la fraction d'argile est de [2, 60] sur la zone couverte par les images analysées contre [5, 22] pour les données d'entraînement.

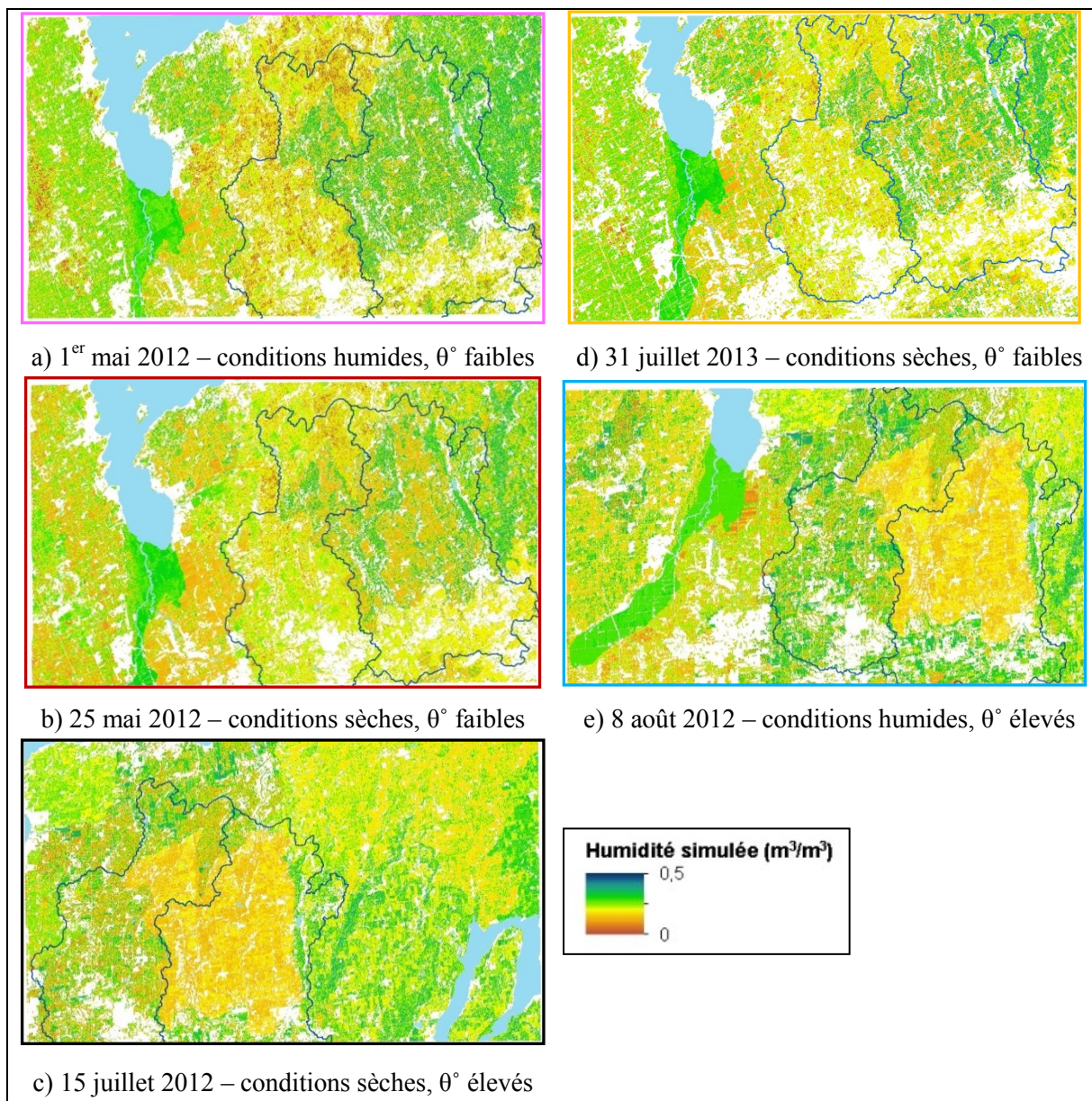


Figure 5.15 Carte d'estimation de l'humidité du sol RNA_{θ° , $\sigma^\circ HH$, % sable, % argile

Compte tenu des résultats obtenus dans cette étude, les fractions de sable et d'argile ne sont pas considérées comme des variables permettant d'améliorer la capacité des RNA à estimer l'humidité du sol. Des cartes de sol plus précises et des mesures de texture aux champs pourraient donner des résultats cartographiques plus probants étant donné les bonnes performances obtenues au niveau des critères de performance.

5.1.4 Angle d'incidence, coefficients de rétrodiffusion HH et HV

Le coefficient de rétrodiffusion en polarisation HV (σ°_{HV}) a été ajouté comme variable entrante aux RNA afin d'évaluer son influence sur la précision et l'incertitude dans les zones où la végétation est abondante.

En terme de performances, le meilleur $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}HH, \sigma^{\circ}HV}$ obtenu (tableau 5.4) est très semblable au $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}HH}$. Bien que le coefficient de Pearson soit plus élevé, cette différence n'est pas significative (p-value > 0,05).

Tableau 5.4 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double

Variables entrantes	Coefficient de Pearson (r)	RMSE (m^3/m^3)	MAE (m^3/m^3)	MRE (%)
θ° et $\sigma^{\circ}HH$	0,647	0,069	0,053	22,2
θ° , $\sigma^{\circ}HH$ et pente	0,639	0,071	0,054	25,5
θ° , $\sigma^{\circ}HH$ et texture	0,723	0,062	0,043	17,8
θ° , $\sigma^{\circ}HH$ et $\sigma^{\circ}HV$	0,693	0,065	0,050	23,0

La pente supérieure de 0,10 par rapport au $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}HH}$ de la régression linéaire (figure 5.16) résulte d'une meilleure estimation des valeurs de m_v supérieures à $0,30 \text{ m}^3/\text{m}^3$ pour lesquelles l'erreur moyenne est de $0,055 \text{ m}^3/\text{m}^3$, comparativement à $0,079 \text{ m}^3/\text{m}^3$ pour le $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}HH}$. Les erreurs sur les valeurs inférieures à $0,30 \text{ m}^3/\text{m}^3$ sont, pour leur part, très similaires.

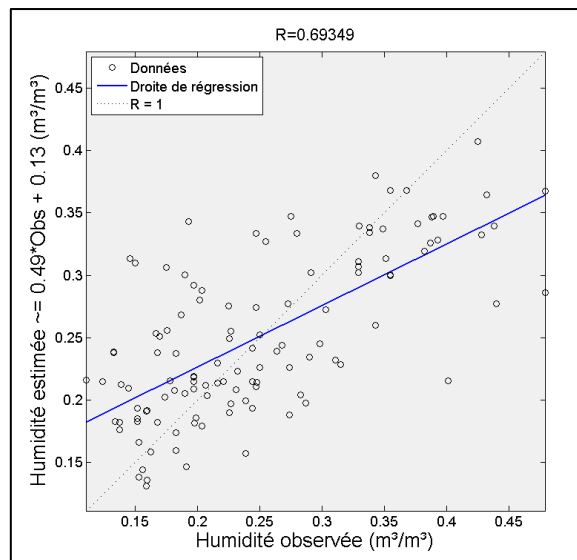


Figure 5.16 Régression linéaire du meilleur $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}HH}$ et $\sigma^{\circ}HV$

Cartes d'estimation de l'humidité et d'incertitude sur l'estimation

Les moyennes des critères de performance des 30 meilleurs $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{HH}}, \sigma^\circ_{\text{HV}}}$ (tableau 5.5) sont semblables à celles pour les $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{HH}}}$, mis à part le coefficient de Pearson qui est légèrement supérieur. Toutefois, les variations plus faibles entre les 30 RNA au niveau de l'erreur relative moyenne (MRE) et, dans une moindre mesure, le coefficient de Pearson, mettent en évidence une stabilité plus grande entre les différents $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{HH}}, \sigma^\circ_{\text{HV}}}$, comparativement aux $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{HH}}}$. En effet, l'écart-type de l'erreur relative moyenne est près de deux fois moins élevé avec 0,7 %, contre 1,3 % pour le $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{HH}}}$. Les différences observées ultérieurement ne peuvent donc pas être attribuées à des différences entre les RNA mais bien à l'influence de l'ajout de σ°_{HV} .

Tableau 5.5 Sommaire des statistiques des 30 meilleurs RNA pour l'approche de polarisation simple et double

Variables entrantes	Moyenne				Domaine de valeur							
	Coefficient de Pearson (r)	RMSE (m^3/m^3)	MAE (m^3/m^3)	MRE (%)	Coefficient de Pearson (r)		RMSE (m^3/m^3)		MAE (m^3/m^3)		MRE (%)	
					Min	Max	Min	Max	Min	Max	Min	Max
θ° et σ°_{HH}	0,592	0,073	0,057	24,7	0,55	0,66	0,068	0,076	0,051	0,061	22,0	26,7
θ° , σ°_{HH} et pente	0,639	0,071	0,054	25,5	0,58	0,70	0,065	0,074	0,050	0,058	21,6	25,1
θ° , σ°_{HH} et texture	0,648	0,069	0,052	21,8	0,61	0,72	0,062	0,074	0,043	0,055	17,8	23,7
θ° , σ°_{HH} et σ°_{HV}	0,627	0,071	0,054	23,6	0,61	0,69	0,065	0,072	0,050	0,056	22,2	24,9

Les cartes obtenues (figure 5.17) présentent de manière générale la même distribution spatiale de l'humidité estimée que celle obtenue avec les $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{HH}}}$. Des différences non négligeables de valeurs d'humidité estimée sont toutefois présentes en fonction de l'angle d'incidence. À angles faibles (figure 5.17 a, b et d), l'humidité estimée a généralement augmenté de plus de $0,05 \text{ m}^3/\text{m}^3$ et de manière plus uniforme dans les zones très mal drainées. L'ajout de σ°_{HV} , qui contient de l'information sur la biomasse via la dépolarisation du signal, semble permettre aux RNA de mieux interpréter l'information provenant de la surface et, donc, de mieux estimer m_v . On observe toutefois une diminution de plus de $0,05 \text{ m}^3/\text{m}^3$ dans les autres zones du bassin versant Pfefferlaw et à l'ouest de celui-ci, pour le 25 mai 2012 et le 31 juillet 2013. Pour les angles élevés (figure 5.17 c et e), l'humidité estimée a diminué de plus de $0,05 \text{ m}^3/\text{m}^3$ principalement dans les zones très mal drainées, à l'exception de certains champs du marais de

Holland (encadré de la figure 5.17 e) où elle a augmenté de manière marquée. Cette augmentation inattendue peut être attribuée au processus d'entraînement des RNA pour lequel la relation établie entre les intrants et l'extrait reste inconnue ainsi qu'à l'écart plus important entre σ°_{HH} et σ°_{HV} dans ces zones.

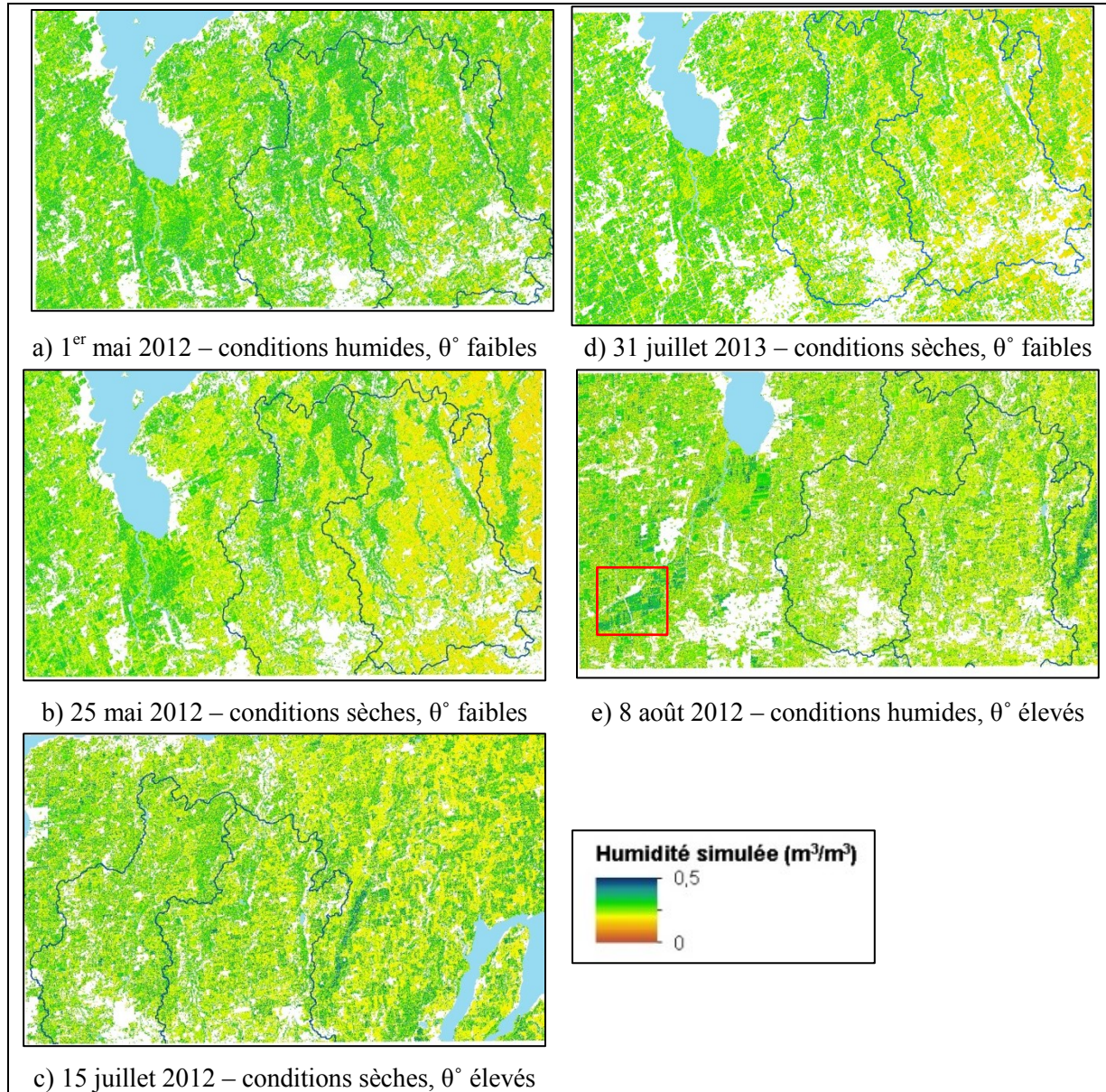


Figure 5.17 Carte d'estimation de l'humidité du sol avec $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}_{HH}, \sigma^{\circ}_{HV}}$

Les histogrammes et les probabilités associées aux différentes classes d'humidité obtenues pour les sous-bassins Baldwin et Pepperlaw sont présentés à la figure 5.18 et le sommaire des variations par rapport aux $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}_{HH}}$ au tableau 5.6. Pour le 1^{er} mai, où l'humidité observée est de $0,42 \text{ m}^3/\text{m}^3$, on remarque principalement une augmentation de $0,13 \text{ m}^3/\text{m}^3$ de la valeur

minimum de m_{ve} par rapport aux $RNA_{\theta^{\circ},\sigma^{\circ}HH}$. Une baisse de 7,6 % de $P(0,10 < x < 0,20)$ de même qu'une hausse de $P(x > 0,30)$ de 17,3 % sont aussi observées. Le 8 août, lorsque la végétation est abondante et que l'humidité est relativement élevée ($0,24 \text{ m}^3/\text{m}^3$), de bons résultats sont aussi obtenus. En effet, les probabilités $P(0,10 < x < 0,20)$ et $P(0,20 < x < 0,30)$ sont respectivement en baisse de 18,6 % et en hausse de 26,8 % par rapport aux $RNA_{\theta^{\circ},\sigma^{\circ}HH}$. Ces résultats semblent donc indiquer une amélioration des estimations en conditions de sols humides indépendamment de l'angle d'incidence malgré une baisse de 8,2 % de $P(x > 0,30)$ le 8 août 2012.

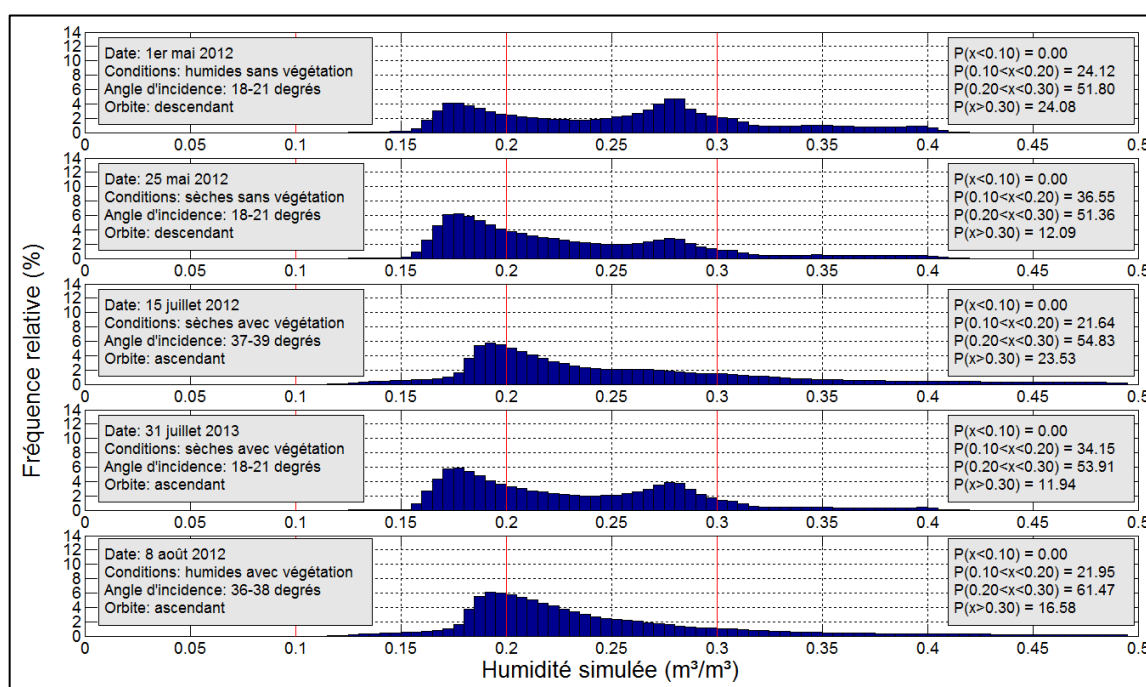


Figure 5.18 Histogrammes d'humidité estimée avec les $RNA_{\theta^{\circ},\sigma^{\circ}HH}$ et $\sigma^{\circ}HV$ pour les sous-bassins versants Baldwin et Pepperlaw

Tableau 5.6 sommaire des variations entre les $RNA_{\theta^{\circ},\sigma^{\circ}HH}$ et $RNA_{\theta^{\circ},\sigma^{\circ}HH,\sigma^{\circ}HV}$

Conditions d'humidité	Quantité de végétation	Date	Variations des probabilités par rapport aux $RNA_{\theta^{\circ},\sigma^{\circ}HH}$ (%), (valeurs $RNA_{\theta^{\circ},\sigma^{\circ}HH,\sigma^{\circ}HV}$ (%))		
			$P(0,10 < x < 0,20)$	$P(0,20 < x < 0,30)$	$P(x > 0,30)$
Humides	Faible	1 ^{er} mai 2012	-7,6 (24,1)	-8,4 (51,8)	+17,3 (51,8)
	Abondante	8 août 2012	-18,6 (22,0)	+26,8 (61,5)	-8,2 (16,6)
Sèches	Faible	25 mai 2012	-11,8 (36,6)	+3,7 (51,36)	+8,2 (12,1)
	Abondante	15 juillet 2012	-15,0 (21,6)	+22,6 (54,8)	-7,4 (23,5)

Par contre, l'ajout de σ°_{HV} semble dégrader les résultats en conditions sèches. En effet, des diminutions de 11,8 % et 15 % de $P(0,10 < x < 0,20)$ ainsi que des hausses de 3,7 % et 22,6 % de $P(0,20 < x < 0,30)$ sont observées, dans l'ordre, pour le 25 mai et le 15 juillet. La probabilité $P(x > 30)$ du 25 mai est aussi plus élevée (8,2 %) alors qu'elle est 7,4 % plus faible le 15 juillet. Cette baisse, bien qu'elle constitue une amélioration, ne contrebalance toutefois pas l'importante augmentation de la probabilité $P(0,20 < x < 0,30)$. Les résultats obtenus pour le 31 juillet sont toutefois très semblables à ceux des $RNA_{\theta^{\circ}, \sigma^{\circ}_{HH}}$, la différence étant principalement une augmentation de 7,1 % de la probabilité $P(x > 30)$, concentrée aux valeurs près de 0,30 m^3/m^3 .

Ces résultats s'expliquent par le fait que σ°_{HV} est très faible pour des sols nus ou faiblement végétalisés et par sa sensibilité élevée à la dépolarisation causée par les réflexions multiples se produisant dans un volume de végétation. Ces deux caractéristiques contribuent à des images de σ°_{HV} , donc à des cartes d'humidité estimée plus contrastées que les images σ°_{HH} . Ces contrastes diminuent toutefois avec la croissance des cultures et une augmentation de l'humidité du sol.

Au niveau de l'incertitude, les statistiques pour les sous-bassins Baldwin et Pefferlaw sont très similaires à celles obtenues avec les $RNA_{\theta^{\circ}, \sigma^{\circ}_{HH}}$ pour les images acquises à faibles angles d'incidence. À angles d'incidence élevés, les incertitudes minimum et maximum sont inférieures. Le nombre de pixels pour lesquels l'incertitude se situe entre 0,05 et 0,10 m^3/m^3 est toutefois plus élevé pour l'ensemble des conditions d'acquisition. La répartition spatiale de l'incertitude est à toute fin pratique identique à celle observée avec θ° et σ°_{HH} comme intrants, et ce, sur l'ensemble des cartes.

L'analyse des performances des RNA ainsi que des cartes obtenues démontrent que l'ajout du coefficient de rétrodiffusion en polarisation HV comme intrant aux réseaux de neurones permet d'améliorer la précision et de réduire l'incertitude de l'humidité estimée. Bien que l'amélioration soit faible et statistiquement non-significative par rapport aux $RNA_{\theta^{\circ}, \sigma^{\circ}_{HH}}$, constat aussi fait par Baghdadi *et al.* [2016], le $RNA_{\theta^{\circ}, \sigma^{\circ}_{HH}, \sigma^{\circ}_{HV}}$ est considéré comme étant le modèle le plus performant et permet l'obtention de cartes d'humidité estimée les plus réalistes parmi celles ayant été développées à l'aide de données RSO en polarisation simple et multiple.

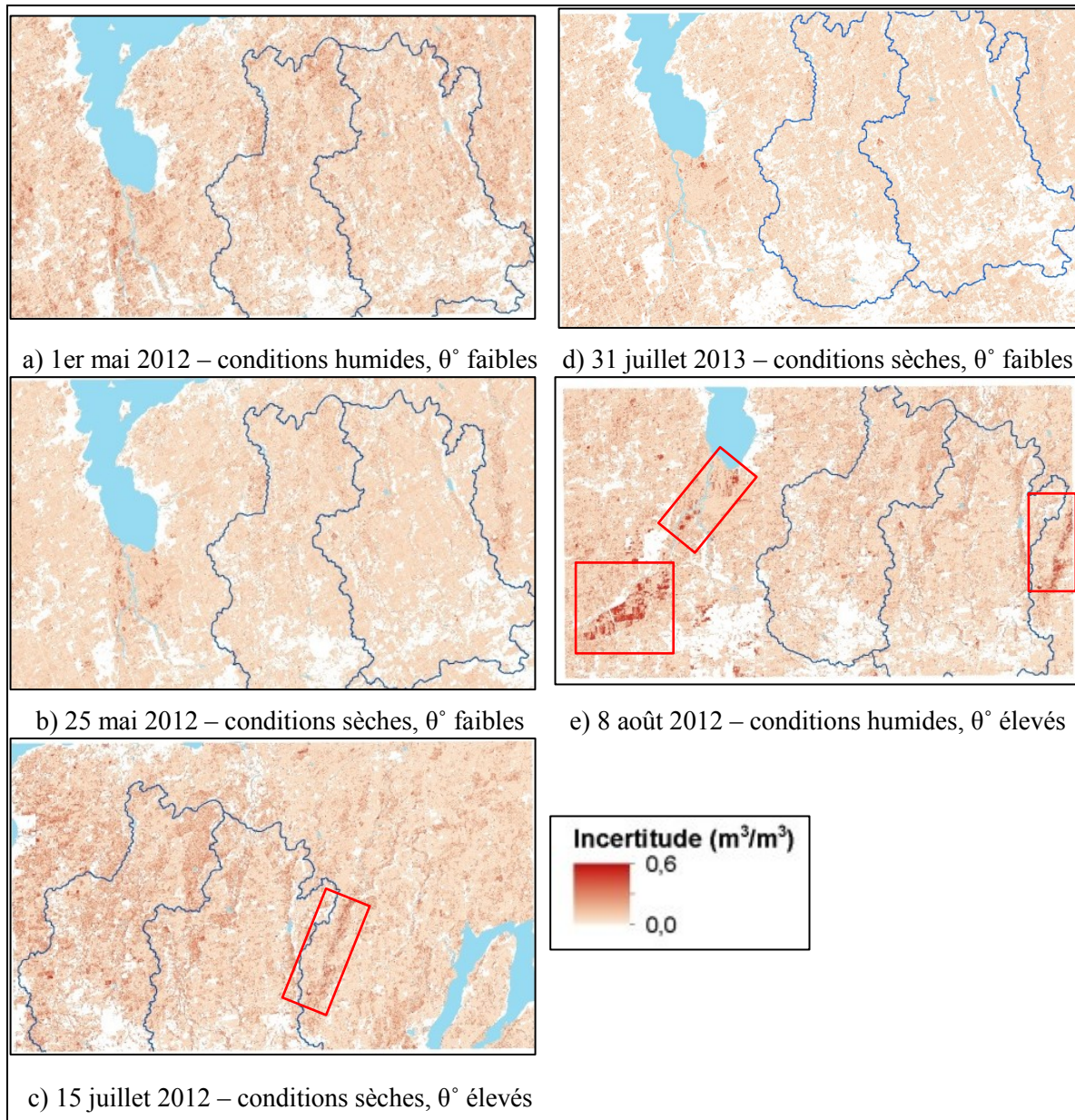


Figure 5.19 Carte d'incertitude sur l'estimation de l'humidité du sol obtenues avec $\text{RNA}_{\theta^\circ, \sigma^\circ\text{HH}} \text{ et } \sigma^\circ\text{HV}$

5.2 Approche polarimétrique

Cette section présente les meilleurs résultats obtenus avec les RNA utilisant un ou plusieurs paramètres polarimétriques dérivés des données RSO acquises en mode polarimétrique, ainsi que leurs analyses. Les résultats de chacune des combinaisons présentées sont comparés avec le meilleur RNA de l'approche simple et double polarisation ($\text{RNA}_{\theta^\circ, \sigma^\circ\text{HH}, \sigma^\circ\text{HV}}$) afin d'établir si la combinaison est mieux adaptée à l'estimation de l'humidité du sol.

5.2.1 Angle d'incidence et puissance de diffusion de surface

Afin de ne considérer que la partie du signal rétrodiffusé provenant de la surface lors de la modélisation de l'humidité du sol, la puissance de diffusion de surface (P_s) a été utilisée avec l'angle d'incidence comme variable entrante lors de l'entraînement des RNA.

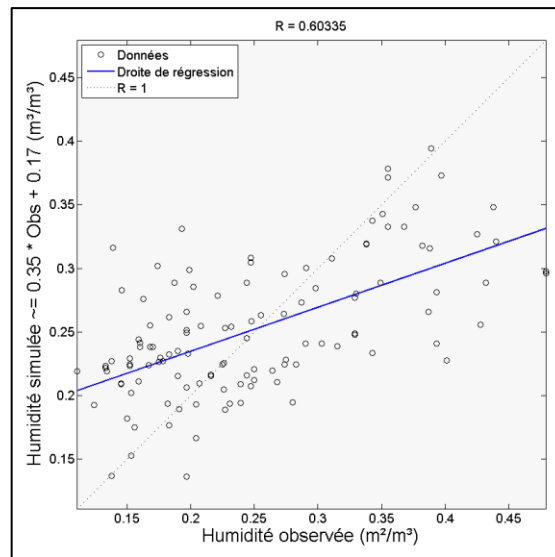
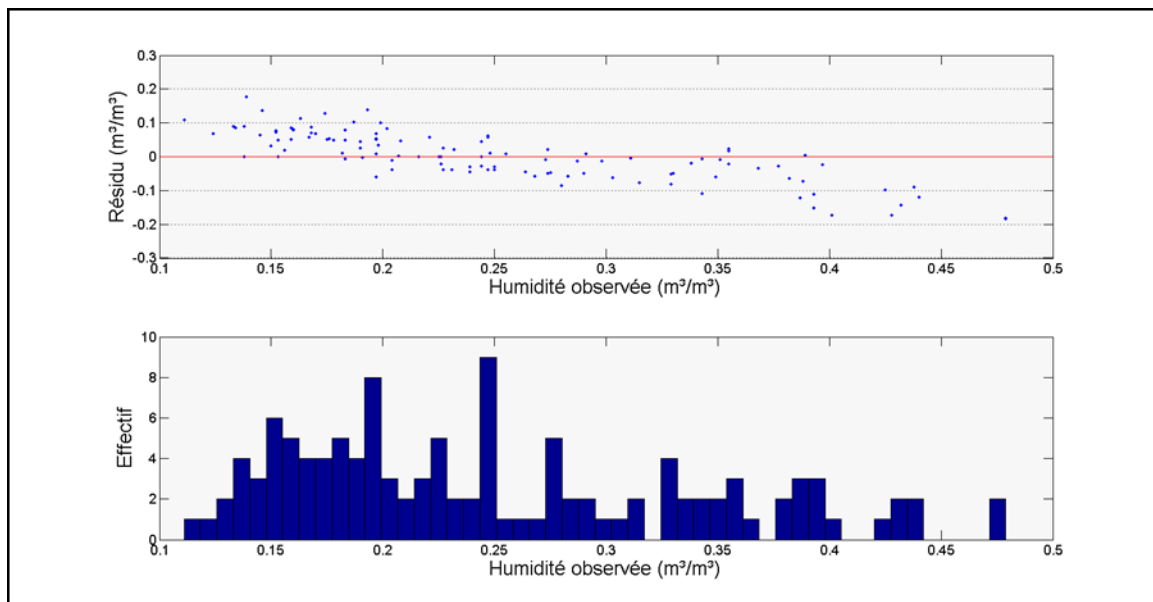
Le meilleur RNA obtenu avec cette combinaison d'intrants possède une couche cachée constituée de 16 neurones. Le tableau 5.7 compare les performances de ce RNA avec celles des RNA obtenus par l'approche en polarisation simple et double, tirées du tableau 5.4. En considérant la dispersion des erreurs et les critères de performance présentés au tableau 5.7, les performances de ce RNA sont inférieures à celles obtenues avec le $\text{RNA}_{\theta^\circ, \sigma^\circ\text{HH}, \sigma^\circ\text{HV}}$, principalement au niveau du coefficient de Pearson et de l'erreur relative moyenne.

Tableau 5.7 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique

Variables entrantes	Approche*	Coefficient de Pearson (r)	RMSE (m^3/m^3)	MAE (m^3/m^3)	MRE (%)
θ° et $\sigma^\circ\text{HH}$	SD	0,647	0,069	0,053	22,2
θ° , $\sigma^\circ\text{HH}$ et pente	SD	0,639	0,071	0,054	25,5
θ° , $\sigma^\circ\text{HH}$ et texture	SD	0,723	0,062	0,043	17,8
θ° , $\sigma^\circ\text{HH}$ et $\sigma^\circ\text{HV}$	SD	0,693	0,065	0,050	23,0
θ° et P_s	P	0,603	0,072	0,057	26,0

* : SD = simple et double; P = polarimétrique

La régression linéaire présentée à la figure 5.20 permet d'observer la relation entre l'humidité observée et celle simulée. La pente de 0,35 met en évidence une sous-estimation des données simulées par rapport aux données observées qui se manifeste particulièrement pour les mesures d'humidité les plus élevées. Comme pour l'ensemble des combinaisons évaluées jusqu'à présent, les résidus sont généralement positifs lorsque l'humidité observée est inférieure à $0,30 \text{ m}^3/\text{m}^3$ et négatifs dans la quasi-totalité des cas au-delà de $0,30 \text{ m}^3/\text{m}^3$ (figure 5.21). Par contre, les erreurs sur les observations supérieures à $0,30 \text{ m}^3/\text{m}^3$ sont plus élevées que pour le $\text{RNA}_{\theta^\circ, \sigma^\circ\text{HH}, \sigma^\circ\text{HV}}$.

Figure 5.20 Régression linéaire du $\text{RNA}_{\theta^{\circ},Ps}$ Figure 5.21 Résidus en fonction de l'humidité du sol (haut) et distribution des observations (bas)) pour le $\text{RNA}_{\theta^{\circ},Ps}$

Cartes d'estimation de l'humidité du sol et de l'incertitude

Les statistiques des 30 meilleurs RNA présentées au tableau 5.8 permettent de constater que les $\text{RNA}_{\theta^{\circ},Ps}$ performant moins bien que les $\text{RNA}_{\theta^{\circ},\sigma^{\circ}HH,\sigma^{\circ}HV}$, principalement au niveau du coefficient de Pearson et de l'erreur relative moyenne.

Tableau 5.8 Sommaire des statistiques des 30 meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique

Variables entrantes	Moyenne				Domaine de valeur							
	Coefficient de Pearson (r)	RMSE (m ³ /m ³)	MAE (m ³ /m ³)	MRE (%)	Coefficient de Pearson (r)		RMSE (m ³ /m ³)		MAE (m ³ /m ³)		MRE (%)	
					Min	Max	Min	Max	Min	Max	Min	Max
θ° et σ_{HH}°	0,592	0,073	0,057	24,7	0,55	0,66	0,068	0,076	0,051	0,061	22,0	26,7
θ° , σ_{HH}° et pente	0,639	0,071	0,054	25,5	0,58	0,70	0,065	0,074	0,050	0,058	21,6	25,1
θ° , σ_{HH}° et texture	0,648	0,069	0,052	21,8	0,61	0,72	0,062	0,074	0,043	0,055	17,8	23,7
θ° , σ_{HH}° et σ_{HV}°	0,627	0,071	0,054	23,6	0,61	0,69	0,065	0,072	0,050	0,056	22,2	24,9
θ° et Ps*	0,555	0,076	0,06	26,8	0,50	0,63	0,070	0,079	0,057	0,063	25,5	28,0

* Approche polarimétrique

Les cartes d'estimation d'humidité du sol (figure 5.22) présentent une variation spatiale de l'humidité du sol estimée en fonction de l'angle d'incidence et de la puissance de diffusion de surface. Les cartes obtenues à angles d'incidence élevés présentent de manière générale une humidité estimée plus faible qu'à angles d'incidence faibles. À angles d'incidence faibles, les estimations sont les plus faibles ($\sim 0,13$ m³/m³) lorsque les valeurs de Ps sont supérieures à -2 dB, ce qui correspond à la limite supérieure du jeu de données d'entraînement. À angles d'incidence élevés, l'humidité du sol estimée augmente avec une augmentation de Ps, ce qui est attendu. Par contre, des valeurs d'humidité du sol estimées se rapprochant de 0,30 m³/m³ sont aussi obtenues pour les valeurs de Ps les plus faibles, qui ne sont pas représentées dans l'échantillon d'entraînement. Les RNA semblent donc sensibles à l'humidité du sol mais incapables d'extrapoler.

Lorsque comparées aux cartes d'estimation d'humidité du sol obtenues avec les RNA $_{\theta^\circ, \sigma_{HH}^\circ, \sigma_{HV}^\circ}$, on observe des différences importantes (figure 5.23). Une réduction supérieure à 0,05 m³/m³ est observée dans les zones très mal drainées pour l'ensemble des cartes. À angles d'incidence faibles (figure 5.23 a, b et d), il y a diminution pour les valeurs de Ps élevées et, inversement, une augmentation de l'humidité pour les faibles valeurs, plus particulièrement pour la carte du 25 mai. À angles d'incidence élevés (figure 5.23 c et e), une diminution supérieure à 0,05 m³/m³ est observée sur la quasi-totalité du territoire et de manière plus importante en conditions sèches (figure 5.23 c).

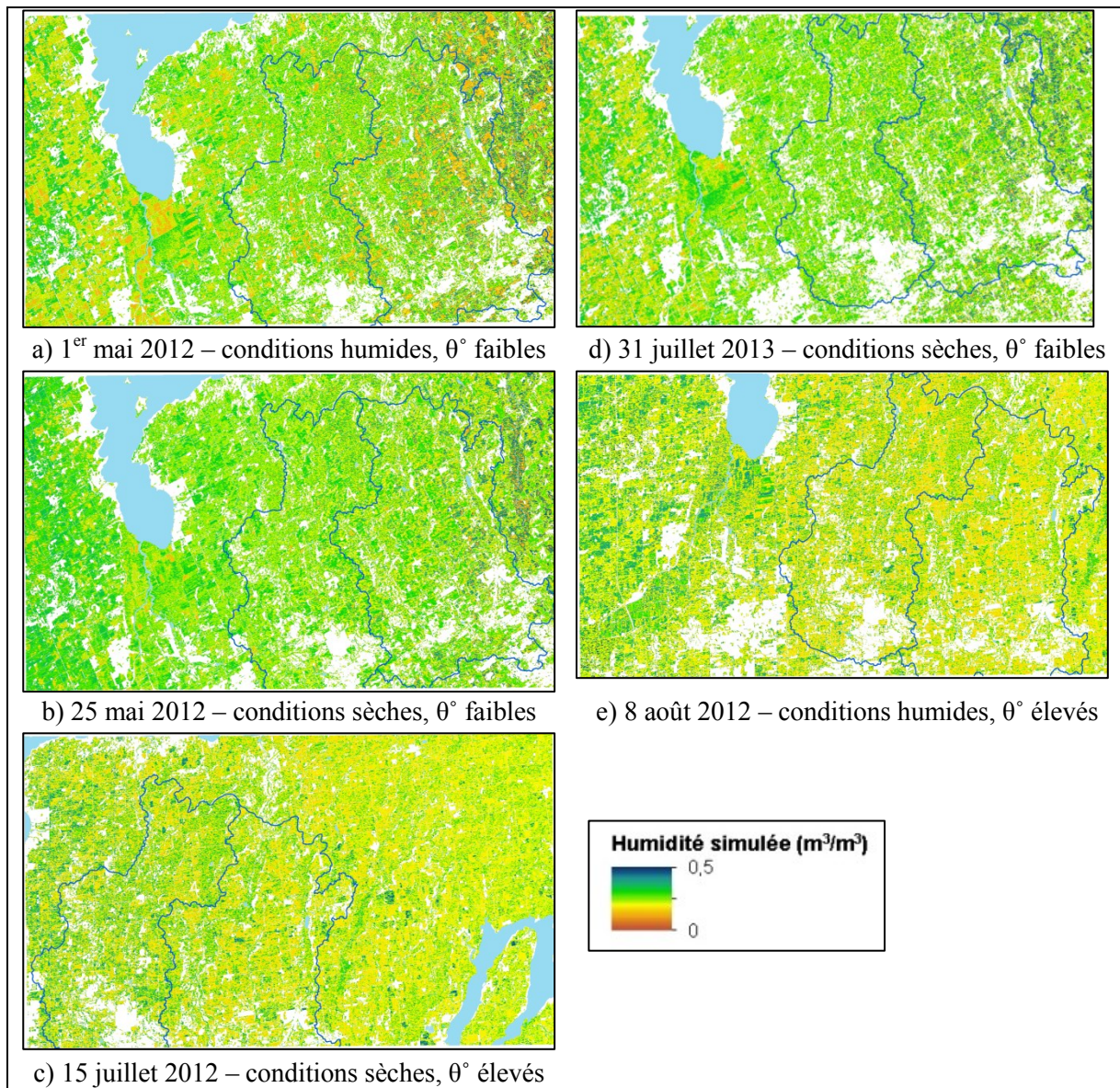


Figure 5.22 Carte d'estimation de l'humidité du sol avec $\text{RNA}_{\theta^\circ, \text{Ps}}$

À angles d'incidence élevés, le signal radar est davantage influencé par la végétation, ce qui résulte en une contribution moins élevée de la surface au signal rétrodiffusé total. Le $\text{RNA}_{\theta^\circ, \text{Ps}}$ aura donc tendance à sous-estimer la quantité d'humidité du sol. En effet, les images du 15 juillet et du 8 août présentent une diffusion de surface plus faible par rapport à celle du 31 juillet 2013 malgré la quantité élevée de végétation dans tous les cas, ce qui met en évidence l'influence de l'angle d'incidence.

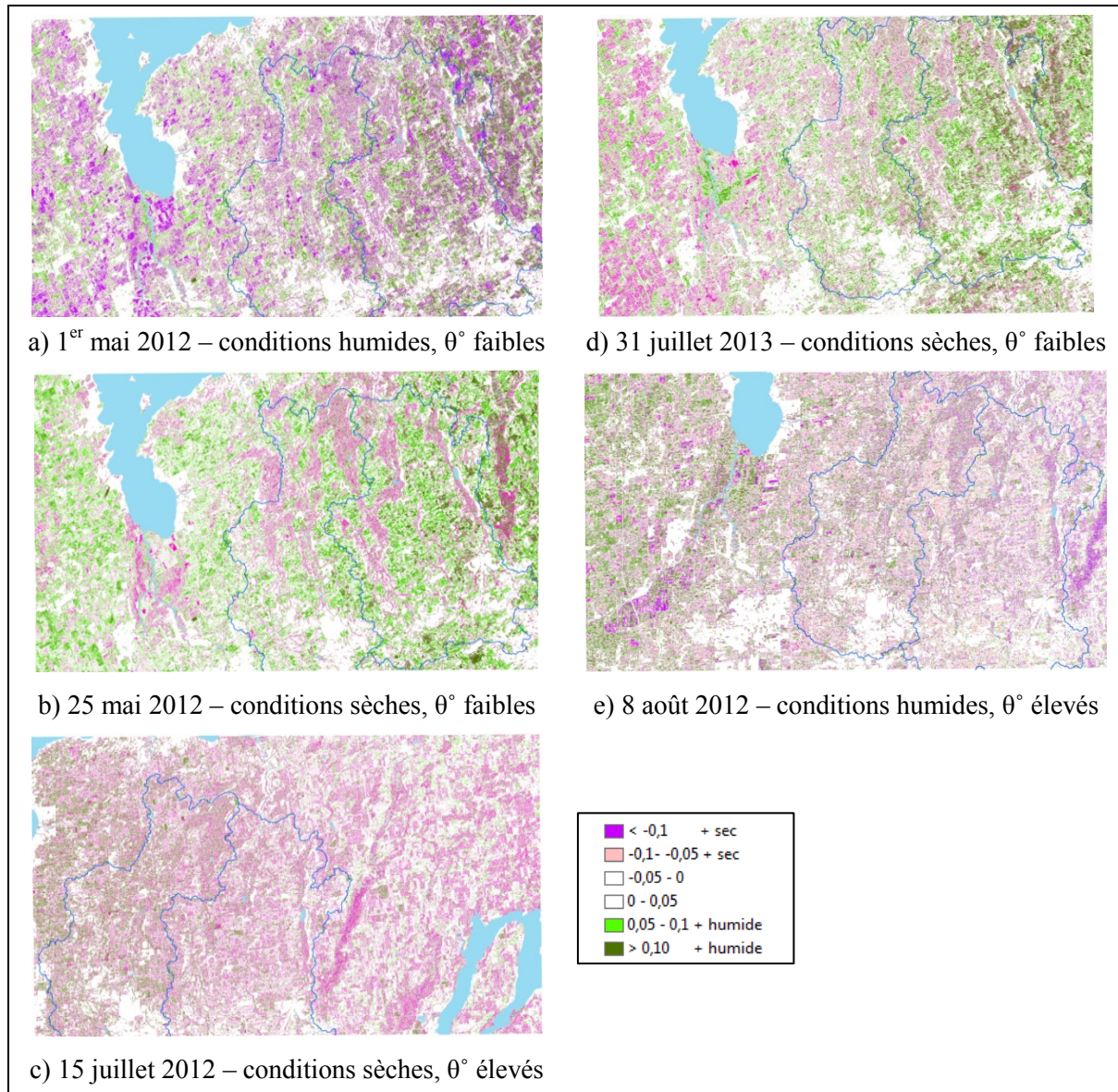


Figure 5.23 Variations d'humidité du sol estimée entre les $RNA_{\theta^{\circ}, \sigma^{\circ}HH, \sigma^{\circ}HV}$ et $RNA_{\theta^{\circ}, P_s}$

L'analyse de la distribution de l'humidité estimée des sous-bassins Baldwin et Pefferlaw (figure 5.24) permet de constater que les $RNA_{\theta^{\circ}, P_s}$ ne sont pas sensibles à l'humidité du sol, tant à angles d'incidence faibles qu'élevés. Tel que prévu, la sensibilité à l'humidité du sol est plus importante à faibles angles d'incidence (1^{er} mai 2012, 25 mai 2012, 31 juillet 2013) qu'à angles d'incidence élevés (15 juillet 2012 et 8 août 2012). À angles d'incidence élevés, les distributions en conditions sèches et humides sont semblables. Les probabilités permettent toutefois d'observer des différences (tableau 5.9). Les probabilités $P(0,10 < x < 0,20)$ et $P(x > 0,30)$ pour les conditions sèches du 15 juillet 2012 sont respectivement inférieures et

supérieures à celles en conditions humides du 8 août 2012. Par contre, à angles d'incidence faibles, les résultats obtenus sont contraires à ce qui est attendu. Pour les conditions de sol humides du 1^{er} mai 2012, où l'humidité observée à la sonde 9915 au moment de l'acquisition de l'image radar est de $0,43 \text{ m}^3/\text{m}^3$, les probabilités $P(x < 0,10)$ et $P(0,10 < x < 0,20)$ sont supérieures, alors que $P(0,20 < x < 0,30)$ est inférieure aux probabilités correspondantes du 25 mai 2012, où l'humidité observée à la sonde est de $0,20 \text{ m}^3/\text{m}^3$.

Tableau 5.9 Sommaire des probabilités à angles d'incidence faibles et en présence de végétation abondante pour les $\text{RNA}_{\theta^\circ \text{Ps}}$

Angles d'incidence	Conditions d'humidité	Quantité de végétation	Date	Probabilités par classe d'humidité estimée (%)			
				$P(x < 0,10)$	$P(0,10 < x < 0,20)$	$P(0,20 < x < 0,30)$	$P(x > 0,30)$
Faibles	Humides	Faible	1 ^{er} mai 2012	2,2	27,7	48,1	22,0
	Sèches	Faible	25 mai 2012	0,4	21,7	58,2	19,8
	Sèches	Abondante	31 juillet 2013	0,5	19,0	58,6	21,9
Élevés	Sèches	Abondante	15 juillet 2012	0,0	44,4	37,0	18,7
	Humides	Abondante	8 août 2012	0,0	52,49	34,9	12,6

Les cartes d'incertitude obtenues présentent une incertitude faible et très homogène pour les pixels pour lesquels la valeur de Ps est dans le domaine de valeurs du jeu de données d'apprentissage, qui est de $[-21,9 \text{ } -2,2 \text{ dB}]$. Pour les pixels ayant des valeurs de Ps supérieures -2 dB , l'incertitude se concentre autour de $0,40 \text{ m}^3/\text{m}^3$ et augmente avec une augmentation de Ps . L'incertitude associée aux quelques pixels ayant des valeurs de Ps inférieures à $-21,9 \text{ dB}$ est du même ordre de grandeur que pour les valeurs de Ps entre $-21,9$ et $-2,2 \text{ dB}$.

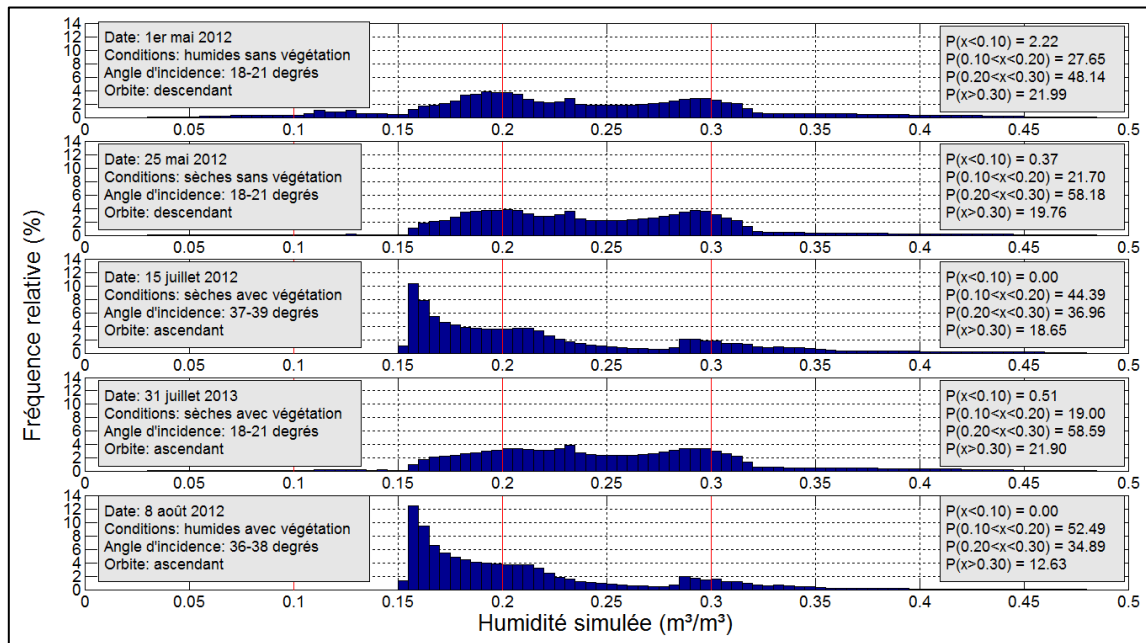


Figure 5.24 Histogrammes de l'humidité estimée avec $RNA_{\theta^{\circ}, P_s}$ pour les sous-bassins versants Baldwin et Pefferlaw

Les résultats obtenus démontrent que l'utilisation de l'angle d'incidence et de la puissance de diffusion de surface comme intrants produit des RNA peu sensibles à l'humidité du sol, et ce, pour toutes les conditions d'acquisition. Il semble donc que l'utilisation de la rétrodiffusion provenant uniquement de la surface ne soit pas suffisante à l'estimation du sol et que la diffusion de volume puisse apporter de l'information nécessaire à celle-ci. Des résultats similaires ont été obtenus par Baghdadi *et al.* [2016] où l'utilisation du paramètre polarimétrique de la diffusion de surface de Pauli n'a pas permis d'améliorer les estimations de m_v .

Les résultats obtenus pour les $RNA_{\theta^{\circ}, P_s}$ dans le cadre de ce projet et par Baghdadi *et al.* [2016] peuvent, entre autres, être dû au fait que les modèles actuels de rétrodiffusion de surface contiennent possiblement des erreurs. Il est aussi probable que les composantes de surface et de volume doivent être prises en compte pour établir correctement la relation entre l'humidité du sol et les puissances de diffusion. Par contre, Hajnsek *et al.* [2009] font l'hypothèse qu'une fois la composante de volume enlevée, les composantes de surface et de double réflexion restantes dépendent des propriétés diélectriques du sol sous-jacent. La principale cause des résultats mitigés obtenus est probablement attribuable au fait d'avoir réalisé l'inversion de m_v sur l'ensemble du territoire. En effet, Hajnsek *et al.* [2009] n'ont réalisé l'inversion que

lorsque que certaines conditions étaient respectées. Ces conditions sont que 1) le mécanisme de diffusion inversé doit être dominant et 2) les paramètres d'inversion doivent être physiquement corrects. Cette procédure a conduit à des taux d'inversion de la diffusion de surface X-Bragg compris entre 1% et 19% en fonction du stade de développement de la végétation. Le taux d'inversion représentant la quantité relative de pixels de l'image pouvant être inversés. Cette approche aurait par contre rendu l'analyse comparative des résultats avec les autres RNA pour le moins difficile.

Par conséquent, la puissance de diffusion de surface n'est pas considérée comme une variable permettant d'améliorer les performances des RNA comparativement à celles obtenues avec les données en polarisation simple et multiple.

5.2.2 Angle d'incidence, puissance de diffusion de surface et de volume

La puissance de diffusion de volume (P_v) a été ajoutée à l'angle d'incidence et à la puissance de diffusion de surface (P_s) comme intrant aux RNA afin de : 1) prendre en compte la diffusion de volume qui, selon les résultats obtenus à la section précédente, semble être nécessaire à l'estimation de l'humidité du sol; 2) améliorer la caractérisation de la diffusion de volume associée à la végétation et l'atténuation de la rétrodiffusion qu'elle cause.

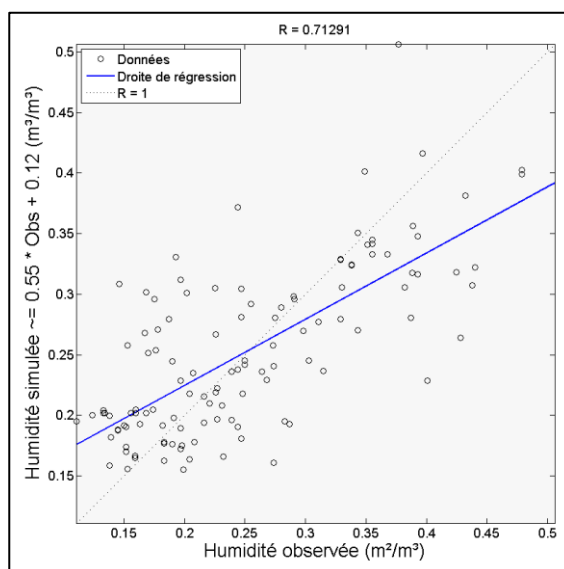
Le meilleur RNA obtenu avec cette combinaison d'intrants possède une couche cachée constituée de 13 neurones. En considérant la dispersion des erreurs et les critères de performance présentés au tableau 5.10, l'ajout de P_v a permis d'améliorer les performances lorsque comparé au RNA_{θ°, P_s} alors que des critères de performance similaires à ceux du $RNA_{\theta^\circ, \sigma^\circ HH, \sigma^\circ HV}$ ont été obtenus. L'erreur relative moyenne est toutefois plus faible avec le $RNA_{\theta^\circ, P_s, P_v}$.

Tableau 5.10 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique

Variables entrantes	Approche *	Coefficient de Pearson (r)	RMSE (m ³ /m ³)	MAE (m ³ /m ³)	MRE (%)
θ° , et $\sigma^\circ\text{HH}$	SD	0,647	0,069	0,053	22,2
θ° , $\sigma^\circ\text{HH}$ et pente	SD	0,639	0,071	0,054	25,5
θ° , $\sigma^\circ\text{HH}$ et texture	SD	0,723	0,062	0,043	17,8
θ° , $\sigma^\circ\text{HH}$ et $\sigma^\circ\text{HV}$	SD	0,693	0,065	0,050	23,0
θ° , et Ps	P	0,603	0,072	0,057	26,0
θ° , Ps et Pv	P	0,713	0,064	0,049	21,9

* : SD = simple et double; P = polarimétrique

La régression linéaire présentée à la figure 5.25 permet d'observer la relation entre l'humidité observée et celle simulée. La pente de 0,55 met en évidence une sous-estimation des données simulées par rapport aux données observées qui se manifeste particulièrement pour les mesures d'humidité les plus élevées.

Figure 5.25 Régression linéaire du $\text{RNA}_{\theta^\circ, \text{Ps}, \text{Pv}}$

Cartes d'estimation de l'humidité du sol et de l'incertitude

Lorsque comparées à celles du meilleur RNA, les statistiques des 30 meilleurs $\text{RNA}_{\theta^\circ, \text{Ps}, \text{Pv}}$ (tableau 5.11) indiquent que les performances sont davantage constantes entre les 30 RNA sélectionnés comparativement aux $\text{RNA}_{\theta^\circ, \sigma^\circ\text{HH}, \sigma^\circ\text{HV}}$. En effet, l'écart entre le coefficient de Pearson moyen des 30 $\text{RNA}_{\theta^\circ, \text{Ps}, \text{Pv}}$ sélectionnés (tableau 5.11) et du meilleur RNA (tableau 5.10) est plus faible comparativement aux $\text{RNA}_{\theta^\circ, \sigma^\circ\text{HH}, \sigma^\circ\text{HV}}$. Le meilleur et les 30 meilleurs

$RNA_{\theta^\circ, \sigma^\circ_{HH}, \sigma^\circ_{HV}}$ possèdent respectivement un r de 0,693 (tableau 5.10) et un r moyen de 0,627 (tableau 5.11), comparativement à 0,713 à 0,672, dans le même ordre, pour les RNA avec θ° , Ps et Pv comme intrants. Un écart moins marqué entre ces mêmes combinaisons d'intrants est observé au niveau des autres critères de performance.

Tableau 5.11 Sommaire des statistiques des 30 meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique

Variables entrantes	Moyenne				Domaine de valeur							
	Coefficient de Pearson (r)	RMSE (m ³ /m ³)	MAE (m ³ /m ³)	MRE (%)	Coefficient de Pearson (r)		RMSE (m ³ /m ³)		MAE (m ³ /m ³)		MRE (%)	
					Min	Max	Min	Max	Min	Max	Min	Max
θ° et σ°_{HH}	0,592	0,073	0,057	24,7	0,55	0,66	0,068	0,076	0,051	0,061	22,0	26,7
θ° , σ°_{HH} et pente	0,639	0,071	0,054	25,5	0,58	0,70	0,065	0,074	0,050	0,058	21,6	25,1
θ° , σ°_{HH} et texture	0,648	0,069	0,052	21,8	0,61	0,72	0,062	0,074	0,043	0,055	17,8	23,7
θ° , σ°_{HH} et σ°_{HV}	0,627	0,071	0,054	23,6	0,61	0,69	0,065	0,072	0,050	0,056	22,2	24,9
θ° et Ps*	0,555	0,076	0,06	26,8	0,50	0,63	0,070	0,079	0,057	0,063	25,5	28,0
θ° , Ps et Pv*	0,672	0,067	0,052	22,8	0,63	0,71	0,064	0,074	0,047	0,057	21,1	24,0

* Approche polarimétrique

La comparaison des cartes d'humidité du sol estimée obtenues avec les $RNA_{\theta^\circ, Ps, Pv}$ et les $RNA_{\theta^\circ, \sigma^\circ_{HH}, \sigma^\circ_{HV}}$ permet d'observer d'importantes différences. De manière générale, la répartition spatiale de m_{ve} est moins homogène et les cartes obtenues à l'aide des $RNA_{\theta^\circ, Ps, Pv}$ présentent une humidité moins élevée à l'exception des sols très mal drainés qui sont plus humides.

Les différences sont les plus marquées pour les cartes représentant les conditions d'humidité du sol les plus sèches et les plus humides (figure 5.26 a, c et e). Une proportion élevée des sols très mal drainés présente une augmentation de m_{ve} supérieure à 0,10 m³/m³ alors que dans le reste de ces zones, l'augmentation se situe entre 0,05 et 0,10 m³/m³. En dehors des zones très mal drainées, on observe principalement une réduction supérieure à 0,10 m³/m³. On observe aussi plusieurs groupements de pixels pour lesquels la diminution dépasse 0,20 m³/m³ sur les cartes du 1^{er} mai et du 8 août 2012 (figure 5.26 a et e).

Les variations suivent les mêmes tendances mais sont moins prononcées en conditions d'humidité moyenne (figure 5.26 b et d). En effet, l'écart se situe majoritairement en deçà de $0,10 \text{ m}^3/\text{m}^3$ à l'extérieur des zones très mal drainées en absence de végétation abondante (figure 5.26 b). En présence de végétation abondante (figure 5.26 d), l'écart se situe majoritairement en deçà de $0,10 \text{ m}^3/\text{m}^3$ sur l'ensemble de la carte.

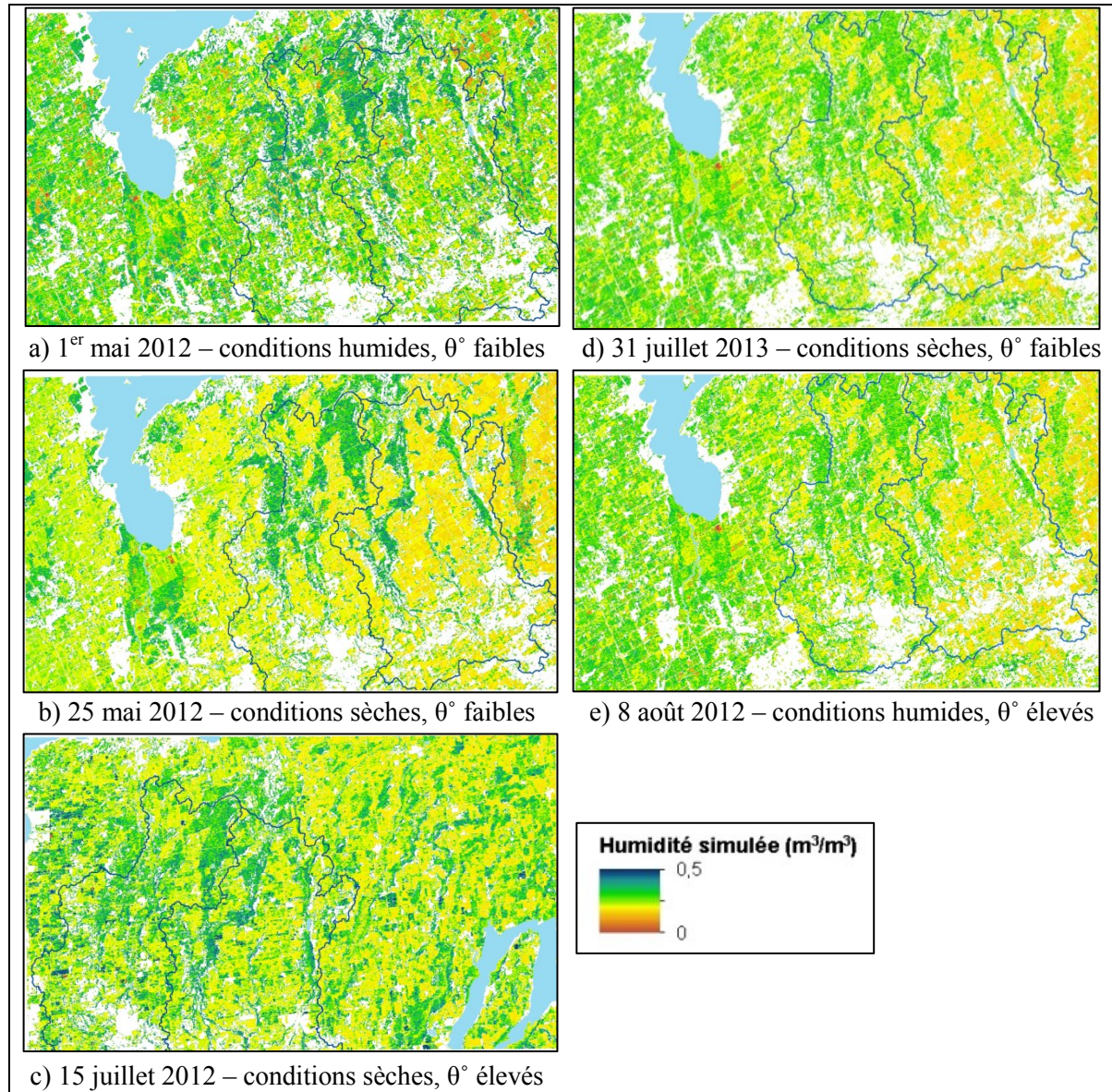


Figure 5.26 Carte d'estimation de l'humidité du sol avec $\text{RNA}_{\theta^\circ, \text{Ps}, \text{Pv}}$

Les cartes d'humidité estimée obtenues présentent généralement les valeurs de m_{ve} parmi les plus élevées dans les zones très mal drainées pour l'ensemble des conditions. À angles d'incidence faibles et en absence de végétation, les $\text{RNA}_{\theta^\circ, \text{Ps}, \text{Pv}}$ démontrent une sensibilité

élevée à l'humidité du sol. Ceci confirme que la diffusion de volume intervient dans la sensibilité du signal radar à l'humidité du sol et appuie la conclusion que l'utilisation de P_s uniquement n'est pas suffisante. La carte du 1^{er} mai 2012 (figure 5.26 a) présente une humidité du sol estimée généralement plus élevée que celle du 25 mai 2012 (figure 5.26 b), ce qui concorde avec les observations d'humidité du sol de 0,43 et 0,20 m^3/m^3 , respectivement, enregistrées à la sonde 9915 au moment de l'acquisition des images radar.

Les histogrammes et les probabilités des sous-bassins versants Baldwin et Pefferlaw de la figure 5.27 appuient cette affirmation. En effet, la probabilité $P(0 < 0,10 < x < 0,20)$ est de 18,26 % en conditions humides contre 43,54 % en conditions sèches alors que $P(x > 0,30)$ est de 30,52 % contre 16,98 %, dans le même ordre.

La présence de végétation abondante diminue toutefois de manière importante la sensibilité à l'humidité du sol. En effet, l'humidité du sol observée les 25 mai 2012 et 31 juillet 2013 est de 0,20 m^3/m^3 alors que la carte du 25 mai 2012 est nettement plus sèche, à l'exception des zones très mal drainées qui sont plus humides (figure 5.26 b et d). Ce qui est mis en évidence par les importantes différences au niveau de la distribution de l'humidité du sol estimée et des probabilités des différentes classes d'humidité du sol des sous-bassins versant Baldwin et Pefferlaw présentés à la figure 5.27.

En comparant les histogrammes et les probabilités pour ces deux dates (25 mai 2012 et 31 juillet 2013), on remarque qu'elles sont très similaires pour les $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}, \sigma^\circ \text{HV}}$ (figure 5.18) et les $\text{RNA}_{\theta^\circ, P_s}$ (figure 5.24). Ces résultats suggèrent que l'ajout de P_v comme variable entrante augmente la sensibilité des RNA à la végétation et diminue celle à l'humidité du sol en présence de végétation abondante par rapport aux $\text{RNA}_{\theta^\circ, P_s}$.

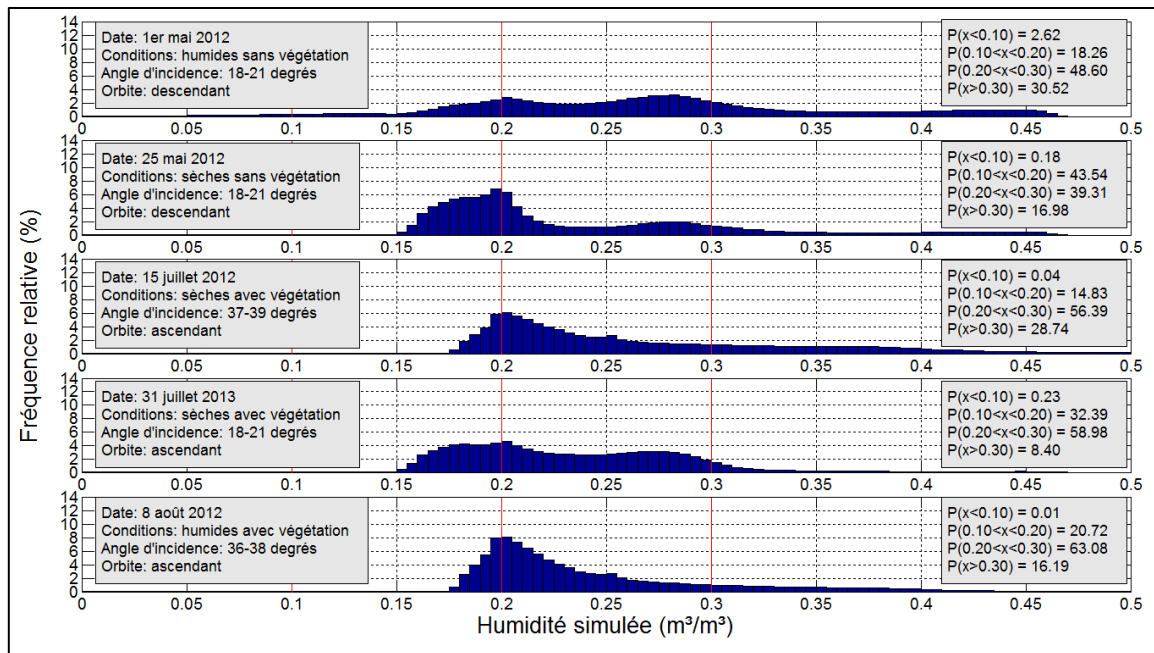


Figure 5.27 Histogrammes de l'humidité estimée avec $RNA_{\theta^{\circ}, P_s, P_v}$ pour les sous-bassins versants Baldwin et Pefferlaw

À angles d'incidence élevés et en présence de végétation abondante, la distribution spatiale de l'humidité du sol estimée dans la zone de recouvrement des deux cartes obtenues (15 juillet et 8 août 2012) est très similaire. La carte obtenue pour les conditions sèches du 15 juillet 2012 (figure 5.26 c) est généralement plus humide que celle pour les conditions humides du 8 août 2012 (figure 5.26 e). En effet, une proportion élevée des pixels de la zone de recouvrement ont des valeurs de m_{ve} supérieures d'au moins $0,05 \text{ m}^3/\text{m}^3$ sur la carte du 15 juillet 2012 par rapport à la carte du 8 août 2012. Bien que l'apparence générale des histogrammes des sous-bassins Baldwin et Pefferlaw soient semblables pour ces deux dates (figure 5.27), les probabilités associées aux différentes classes d'humidité du sol estimée sont fort différentes (tableau 5.12). En effet, la probabilité $P(0 < 0,10 < x < 0,20)$ est de 14,8 % en conditions sèches (15 juillet 2012) comparativement à 20,7 % en conditions humides (8 août) alors que la probabilité $P(x > 0,30)$ est de 28,7 % en conditions sèches et de 16,2 % en conditions humides. Ces résultats contradictoires démontrent une faible sensibilité à l'humidité du sol en conditions de sol sec et de végétation abondante. Ces résultats illustrent également que ce n'est pas parce qu'un RNA produit de meilleures statistiques (ex : RMSE, coefficient de Pearson) que les humidités du sol estimées sont plus représentatives de la réalité. Pour aider

à comprendre davantage ces résultats, les composantes P_s , P_v et P_d du signal total ont été scrutées.

Tableau 5.12 Sommaire des probabilités à angles d'incidence élevés et en présence de végétation abondante pour les $RNA_{\theta^{\circ}, P_s, P_v}$

Conditions d'humidité	Date	Probabilités par classe d'humidité estimée (%)	
		$P(0,10 < x < 0,20)$	$P(x > 0,30)$
Sèches	15 juillet 2012	14,8	28,7
Humides	8 août 2012	20,7	16,2

Les résultats obtenus avec les $RNA_{\theta^{\circ}, P_s, P_v}$ s'expliquent par les contributions de la diffusion de surface et de volume à la puissance de diffusion totale (SPAN) ainsi que le contenu en eau et la nature de la cible. En effet, le contenu en eau de la cible influence la magnitude de SPAN alors que sa nature influence la partition entre les composantes P_s , P_v et P_d lors de la décomposition de la cible. La contribution de P_d étant minime sur l'ensemble de la superficie couverte par les images analysées implique qu'une augmentation de la contribution de P_s entraîne une diminution de celle de P_v . Ainsi, en début de saison les cultures peu développées génèrent principalement de la diffusion de surface, et à mesure que la végétation se développe, celle-ci diminue au profit d'une augmentation de la diffusion de volume. Pour leur part, les autres végétaux, tel que les forêts et les marécages, produisent principalement de la diffusion de volume tout au long de l'année. Les contrastes davantage marqués en début de saison sur les images de décomposition de la cible composites (figure 5.28) mettent en évidence ces variations en fonction du stade de développement de la végétation.

On peut conclure à partir de ces images que les $RNA_{\theta^{\circ}, P_s, P_v}$ sont plus sensibles à la puissance de diffusion de volume que celle de surface. En effet, les zones pour lesquelles la diffusion de volume est dominante sont systématiquement plus humides que celles où la diffusion de surface domine. L'humidité du sol estimée augmente avec la puissance de diffusion, indépendamment du mécanisme de diffusion dominant. En effet, l'humidité élevée du 1^{er} mai 2012 a généré des diffusions de surface et de volume élevées ce qui s'est traduit par la carte la plus humide. Les estimations de l'humidité du sol associées aux valeurs supérieures au domaine de valeurs de P_s du jeu de données d'entraînement (-2 dB) sont toutefois inférieures à $0,10 \text{ m}^3/\text{m}^3$. Cela met en évidence l'importance d'avoir un jeu de données d'entraînement représentatif des conditions sur lesquelles les RNA seront appliqués.

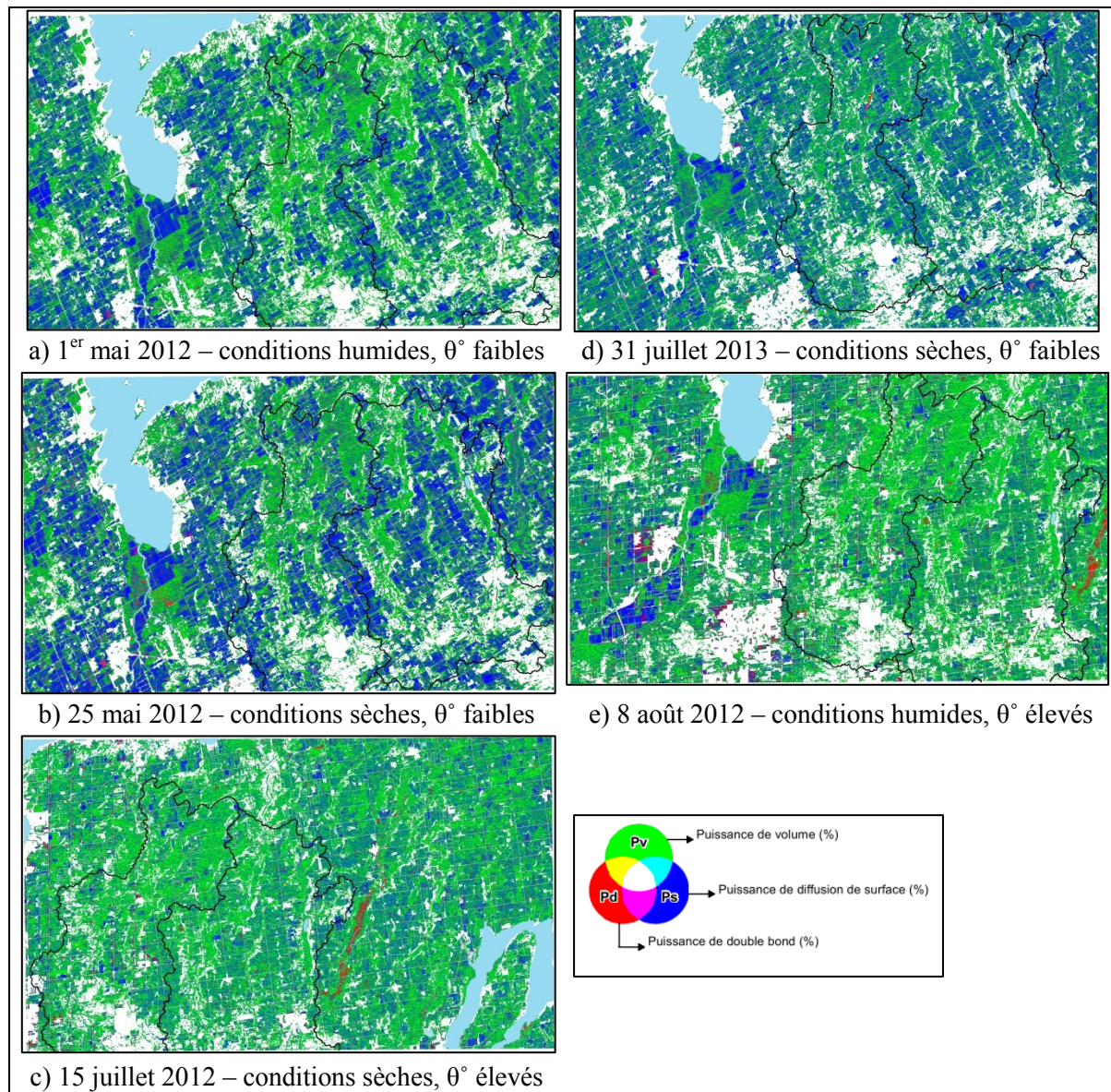


Figure 5.28 Images de décomposition polarimétrique composé RVB – P_d ; P_v ; P_s

L'humidité du sol élevée du 1^{er} mai 2012 a généré de nombreux pixels avec des valeurs de P_s supérieures à -2 dB, ce qui explique la probabilité $P(x < 10)$ de 2,62 % ainsi que les zones nettement plus sèches de la figure 5.26 a. Des valeurs de P_s supérieures à -2 dB ont aussi été mesurées dans la marais de Holland le 8 août 2012 avec le même résultat. Dans ce cas-ci on peut émettre l'hypothèse que comme il s'agit d'une zone maraîchère, les récoltes avaient déjà été faites lors de l'acquisition de l'image. Le fait que la carte du 31 juillet 2013 soit en général plus humide que celle du 25 mai 2012, alors que m_{v0} est de 0,20 m³/m³ dans les deux cas, s'explique elle aussi par la sensibilité plus élevée des RNA à P_v . En effet, l'image de P_v du 31

juillet 2013 se caractérise par une majorité de pixels ayant des valeurs de P_v élevées en raison d'une végétation plus développée, auxquelles des valeurs de m_v élevées sont associées. D'autre part, les images de P_s sont similaires pour les deux dates de même que l'humidité du sol estimée pour les pixels où la diffusion de surface est dominante pour les deux dates.

L'incertitude sur l'humidité estimée obtenue avec les $RNA_{\theta^\circ, P_s, P_v}$ (figure 5.29) est généralement plus faible que celle obtenue avec les $RNA_{\theta^\circ, \sigma^\circ_{HH}, \sigma^\circ_{HV}}$, bien que leur répartition spatiale soit très semblable.

À angles d'incidence faibles (figure 5.29 a, b et d), la répartition de l'incertitude est davantage homogène qu'à angles d'incidence élevés. L'incertitude est maximale pour les valeurs de P_s supérieures au domaine de valeurs de l'échantillon d'entraînement, et dans une moindre mesure P_v . À angles d'incidence élevés, l'incertitude est fonction tant des valeurs élevées de P_s et P_v et, dans une moindre mesure, les faibles valeurs. En effet, l'incertitude est maximale pour les valeurs supérieures à -2 dB et -3 dB pour P_s et P_v , respectivement. La diminution de l'incertitude par rapport aux $RNA_{\theta^\circ, \sigma^\circ_{HH}, \sigma^\circ_{HV}}$ est aussi la plus marquée pour les cartes obtenues à partir des images acquises à angles d'incidence élevés.

L'analyse statistique ne permet pas d'affirmer que l'utilisation des puissances de diffusion de surface et de volume comme intrants permette d'améliorer les performances de manière significative. De plus, les variations entre les performances des 30 RNA sélectionnés sont semblables à celles des $RNA_{\theta^\circ, \sigma^\circ_{HH}, \sigma^\circ_{HV}}$ (tableau 5.11). L'analyse des cartes d'humidité estimée, des histogrammes ainsi que des probabilités indique que les $RNA_{\theta^\circ, P_s, P_v}$ sont moins sensibles à l'humidité du sol comparativement aux $RNA_{\theta^\circ, \sigma^\circ_{HH}, \sigma^\circ_{HV}}$, suggérant que la physique derrière la relation entre m_v et P_s & P_v ne fonctionne pas. Par conséquent, les $RNA_{\theta^\circ, \sigma^\circ_{HH}, \sigma^\circ_{HV}}$ sont considérés comme plus performants et plus robustes car respectant davantage la physique du phénomène de rétrodiffusion en cause. Les images de décomposition de la cible pourraient toutefois être utilisées afin de masquer les pixels pour lesquels la proportion du signal rétrodiffusé provient principalement de la diffusion de volume, donc de la végétation.

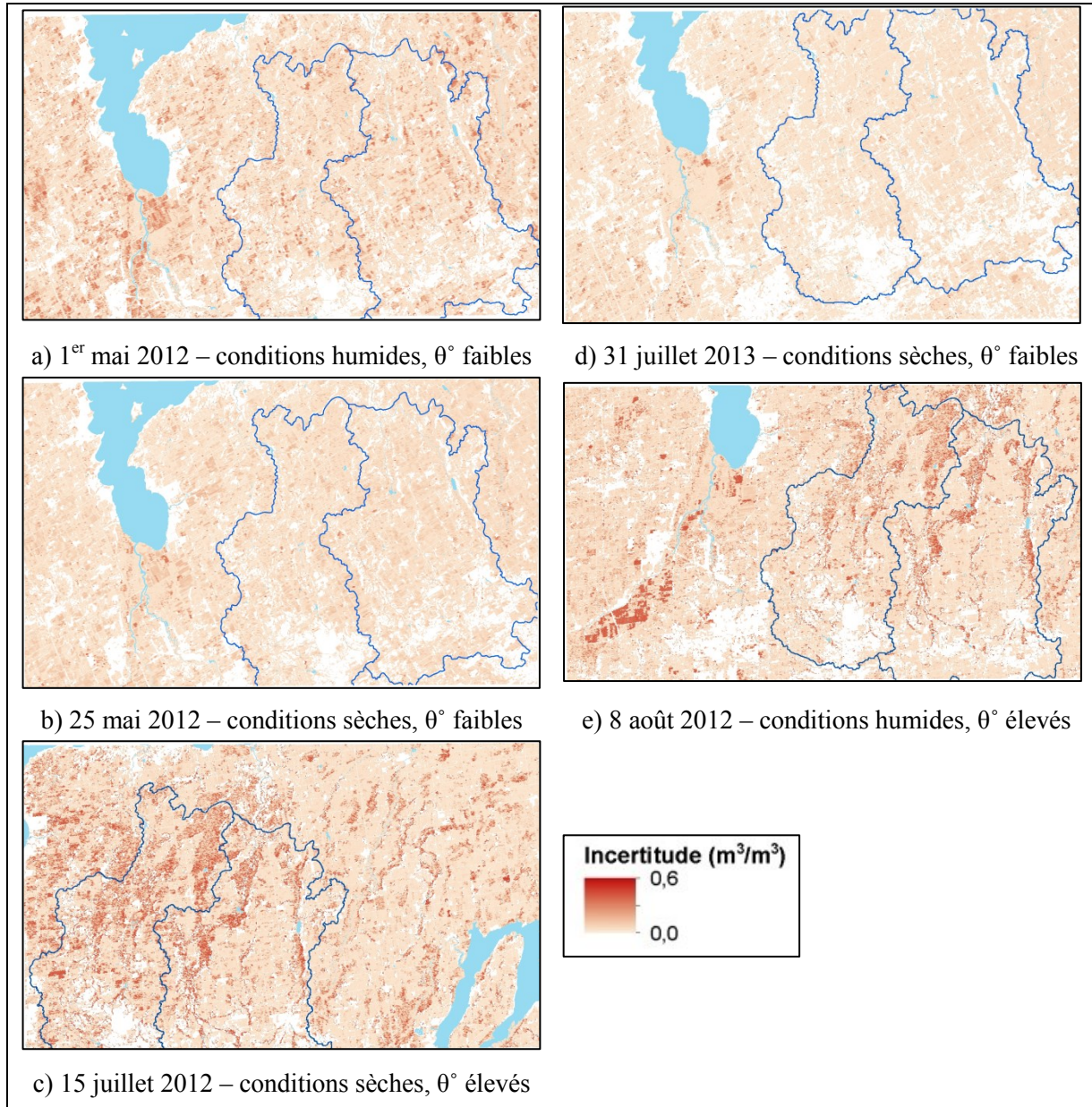


Figure 5.29 Carte d'incertitude sur l'estimation de l'humidité du sol obtenues avec $\text{RNA}_{\theta^\circ, \text{Ps}, \text{Pv}}$

5.2.3 Angle d'incidence, coefficient de rétrodiffusion de surface HH

Dans le but d'éliminer les contributions de la diffusion de volume et de double réflexion au signal radar rétrodiffusé et ainsi conserver uniquement la diffusion provenant de la surface, le coefficient de rétrodiffusion de surface en polarisation HH ($\sigma^\circ_{\text{Surf_HH}}$) a été ajouté à l'angle d'incidence comme variable entrante lors de l'entraînement des RNA. L'utilisation de $\sigma^\circ_{\text{Surf_HH}}$ est aussi motivée par les bons résultats obtenus avec les $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{HH}}}$.

Le meilleur RNA obtenu avec cette combinaison d'intrants possède une couche cachée constituée de 14 neurones. Au niveau statistique, les performances obtenues (tableau 5.13) sont les moins bonnes des combinaisons d'intrants évaluées, et ce, pour l'ensemble des critères. L'utilisation de la diffusion provenant uniquement de la surface a résulté en une dégradation des performances lorsque comparées à celles obtenues avec la rétrodiffusion totale en polarisation HH. En effet, les performances du $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}}$ sont supérieures, et ce, principalement au niveau du coefficient de corrélation (r) et de l'erreur relative moyenne (MRE). Les différences sont encore plus importantes lorsque comparées au $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}, \sigma^\circ \text{HV}}$.

Tableau 5.13 Sommaire des performances des meilleurs RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique

Variables entrantes	Approche*	Coefficient de Pearson (r)	RMSE (m^3/m^3)	MAE (m^3/m^3)	MRE (%)
θ° et $\sigma^\circ \text{HH}$	SD	0,647	0,069	0,053	22,2
θ° , $\sigma^\circ \text{HH}$ et pente	SD	0,639	0,071	0,054	25,5
θ° , $\sigma^\circ \text{HH}$ et texture	SD	0,723	0,062	0,043	17,8
θ° , $\sigma^\circ \text{HH}$ et $\sigma^\circ \text{HV}$	SD	0,693	0,065	0,050	23,0
θ° et Ps	P	0,603	0,072	0,057	26,0
θ° , Ps et Pv	P	0,713	0,064	0,049	21,9
θ° et $\sigma^\circ \text{Surf_HH}$	P	0,563	0,075	0,061	27,0

* : SD = simple et double; P = polarimétrique

La régression linéaire présentée à la figure 5.30 permet d'observer la relation entre l'humidité observée et celle simulée. La pente de 0,38 met en évidence une sous-estimation des données simulées par rapport aux données observées qui se manifeste particulièrement pour les mesures d'humidité les plus élevées. L'ordonnée à l'origine de 0,16 indique pour sa part une surestimation des faibles valeurs observées. Ces résultats sont très semblables à ce qui a été obtenu avec le $\text{RNA}_{\theta^\circ, \sigma^\circ \text{HH}}$ pour lequel l'équation de la droite de régression était $m_{\text{ve}} = 0,39 * m_{\text{vo}} + 0,15$. Les performances inférieures du $\text{RNA}_{\theta^\circ, \sigma^\circ \text{Surf_HH}}$ sont donc uniquement attribuables aux erreurs plus importantes lors de la phase d'entraînement. Il faut toutefois rester prudent sur l'interprétation des performances en se basant uniquement sur les critères de performance. En effet, il a été observé précédemment que de meilleures performances en termes statistiques ne sont pas garantes de l'obtention de cartes d'humidité estimée réalistes.

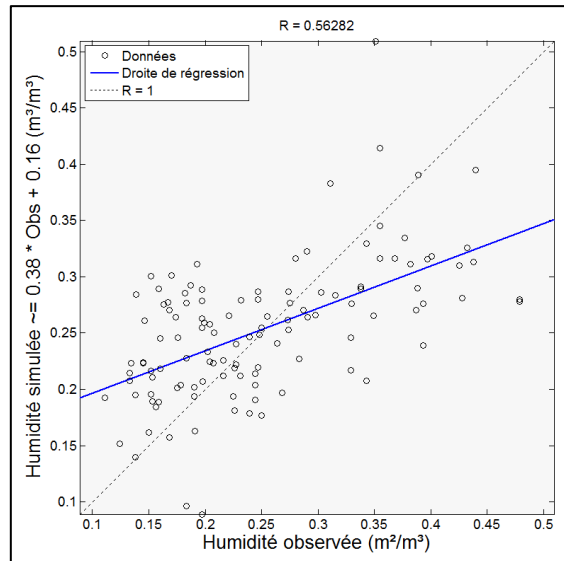


Figure 5.30 Régression linéaire du meilleur $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{Surf_HH}}}$

La relation établie entre la rétrodiffusion radar et l'humidité observée lors de l'apprentissage, représentée par les nuages de points de la figure 5.31, démontre une faible sensibilité de $\sigma^\circ_{\text{Surf_HH}}$ à l'humidité, et ce, indépendamment de l'angle d'incidence. En effet, une augmentation de l'intensité de la rétrodiffusion n'est pas associée à une augmentation de l'humidité du sol observée, contrairement à σ°_{HH} (figure 5.32). Les mauvaises performances ainsi que la faible sensibilité de $\sigma^\circ_{\text{Surf_HH}}$ à l'humidité du sol laissent présager que les cartes d'humidité estimée obtenues avec les $\text{RNA}_{\theta^\circ, \sigma^\circ_{\text{Surf_HH}}}$ sont peu fiables.

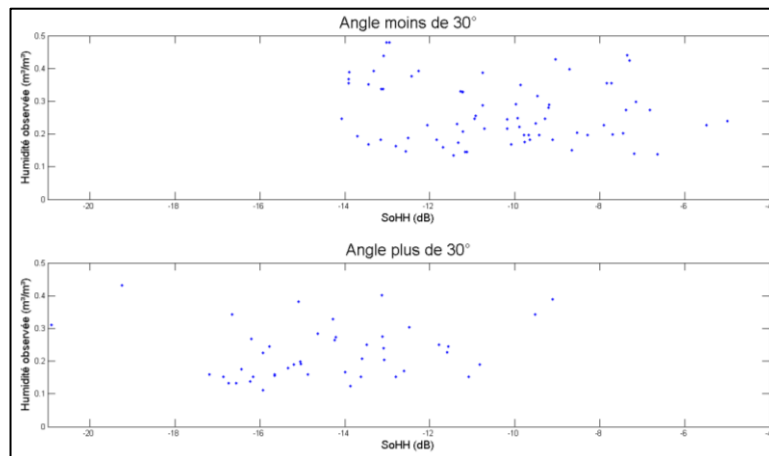


Figure 5.31 Humidité du sol observée par rapport à $\sigma^\circ_{\text{Surf_HH}}$ en fonction de l'angle d'incidence

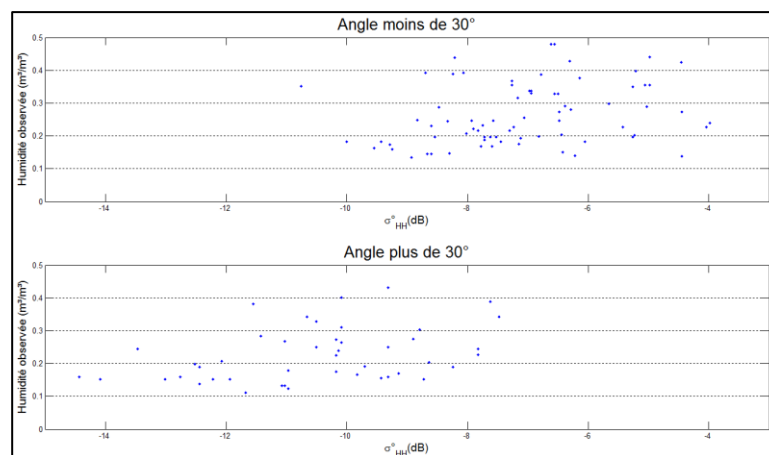


Figure 5.32 Humidité du sol observée par rapport à σ°_{HH} en fonction de l'angle d'incidence

Cartes d'estimation de l'humidité du sol et de l'incertitude

Les statistiques des 30 meilleurs $RNA_{\theta^{\circ}, \sigma^{\circ}_{Surf_HH}}$ présentées au tableau 5.14 vont dans le même sens que ce qui a été observé pour le meilleur des 30 RNA : cette combinaison d'intrants offre les moins bonnes performances parmi les combinaisons évaluées.

Tableau 5.14 Sommaire des statistiques des 30 meilleurs RNA RNA pour l'approche de polarisation simple et double et pour l'approche polarimétrique

Variables entrantes	Moyenne				Domaine de valeur							
	Coefficient de Pearson (r)	RMSE (m³/m³)	MAE (m³/m³)	MRE (%)	Coefficient de Pearson (r)		RMSE (m³/m³)		MAE (m³/m³)		MRE (%)	
					Min	Max	Min	Max	Min	Max	Min	Max
θ° et σ°_{HH}	0,592	0,073	0,057	24,7	0,55	0,66	0,068	0,076	0,051	0,061	22,0	26,7
θ° , σ°_{HH} et pente	0,639	0,071	0,054	25,5	0,58	0,70	0,065	0,074	0,050	0,058	21,6	25,1
θ° , σ°_{HH} et texture	0,648	0,069	0,052	21,8	0,61	0,72	0,062	0,074	0,043	0,055	17,8	23,7
θ° , σ°_{HH} et σ°_{HV}	0,627	0,071	0,054	23,6	0,61	0,69	0,065	0,072	0,050	0,056	22,2	24,9
θ° et Ps^*	0,555	0,076	0,06	26,8	0,50	0,63	0,070	0,079	0,057	0,063	25,5	28,0
θ° , Ps et P_V^*	0,672	0,067	0,052	22,8	0,63	0,71	0,064	0,074	0,047	0,057	21,1	24,0
θ° et $\sigma^{\circ}_{Surf_HH}^*$	0,49	0,079	0,063	27,8	0,44	0,56	0,075	0,082	0,060	0,067	26,8	29,1

* Approche polarimétrique

De manière générale, les cartes d'humidité du sol estimée obtenues à partir des images acquises à angles d'incidence faibles (figure 5.33 a, b et d) présentent des distributions et des valeurs d'humidité du sol similaires entre elles. Cette observation concorde avec la faible

sensibilité de $\sigma^{\circ}_{\text{Surf_HH}}$ à l'humidité observée à la figure 5.31. Les $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}_{\text{Surf_HH}}}$ seraient aussi peu sensibles à la présence de végétation abondante puisque pour une même humidité du sol mesurée à différents stades de croissance de la végétation, des cartes similaires ont été obtenues.

Les cartes obtenues à partir des images acquises à des angles d'incidence élevées (figure 5.33 c et e) sont généralement plus sèches que celles à angles d'incidence faibles. La carte du 8 août 2012 (figure 5.33 e) est la plus sèche malgré une humidité observée de $0,24 \text{ m}^3/\text{m}^3$ contre $0,13 \text{ m}^3/\text{m}^3$ pour le 15 juillet. Ce résultat était attendu compte tenu de l'insensibilité de $\sigma^{\circ}_{\text{Surf_HH}}$ à l'humidité du sol mise en évidence à la figure 5.31.

Les histogrammes et les classes de probabilités des sous-bassins versants Baldwin et Pepperlaw présentés à la figure 5.34 et le sommaire du tableau 5.15 appuient les résultats observés sur les cartes. En effet, les distributions de l'humidité du sol estimée des cartes obtenues à partir des images acquises à faibles angles d'incidence sont très similaires entre elles, de même qu'entre celles acquises à angles d'incidence élevés. L'absence d'influence de la végétation abondante sur l'humidité estimée est mise en évidence par des distributions et des probabilités pratiquement identiques pour le 25 mai 2012 et du 31 juillet 2013, jours pour lesquels m_{v0} est de $0,20 \text{ m}^3/\text{m}^3$.

Tableau 5.15 Sommaire des probabilités à angles d'incidence faibles et en présence de végétation abondante pour les $\text{RNA}_{\theta^{\circ} \text{Ps}, \text{Pv}}$

Angles d'incidence	Conditions d'humidité	Quantité de végétation	Date	Probabilités par classe d'humidité estimée (%)		
				$P(x < 0,10)$	$P(0,10 < x < 0,20)$	$P(x > 0,30)$
Faibles	Humides	Faible	1 ^{er} mai 2012	4,7	9,8	27,0
	Sèches	Faible	25 mai 2012	0,9	1,7	25,3
	Sèches	Abondante	31 juillet 2013	1,3	2,2	24,4
Élevés	Sèches	Abondante	15 juillet 2012	0,0	37,4	18,3
	Humides	Abondante	8 août 2012	0,0	46,1	9,7

La comparaison des résultats obtenus pour le 1^{er} mai 2012 avec ceux du 25 mai 2012 et du 31 juillet 2013 met en évidence la faible sensibilité des RNA à l'humidité du sol observée à angles d'incidence faibles. En effet, malgré une humidité du sol observée élevée ($0,43 \text{ m}^3/\text{m}^3$) et l'importante rétrodiffusion de surface qui en découle, les probabilités $P(x < 0,10)$ et

$P(0,10 < x < 0,20)$ du 1^{er} mai 2012 sont les plus élevées alors que les probabilités $P(x > 0,30)$ sont du même ordre de grandeur pour les trois dates. Les histogrammes obtenus à angles d'incidence élevés présentent aussi des distributions semblables. Les statistiques présentent toutefois des différences contraires au comportement attendu. En effet, pour les conditions humides du 8 août 2012 les probabilités $P(0,10 < x < 0,20)$ et $P(x > 0,30)$ sont de 46,1 % et 9,7 %, respectivement, contre 37,4 % et 18,3 % pour les conditions sèches du 15 juillet 2012.

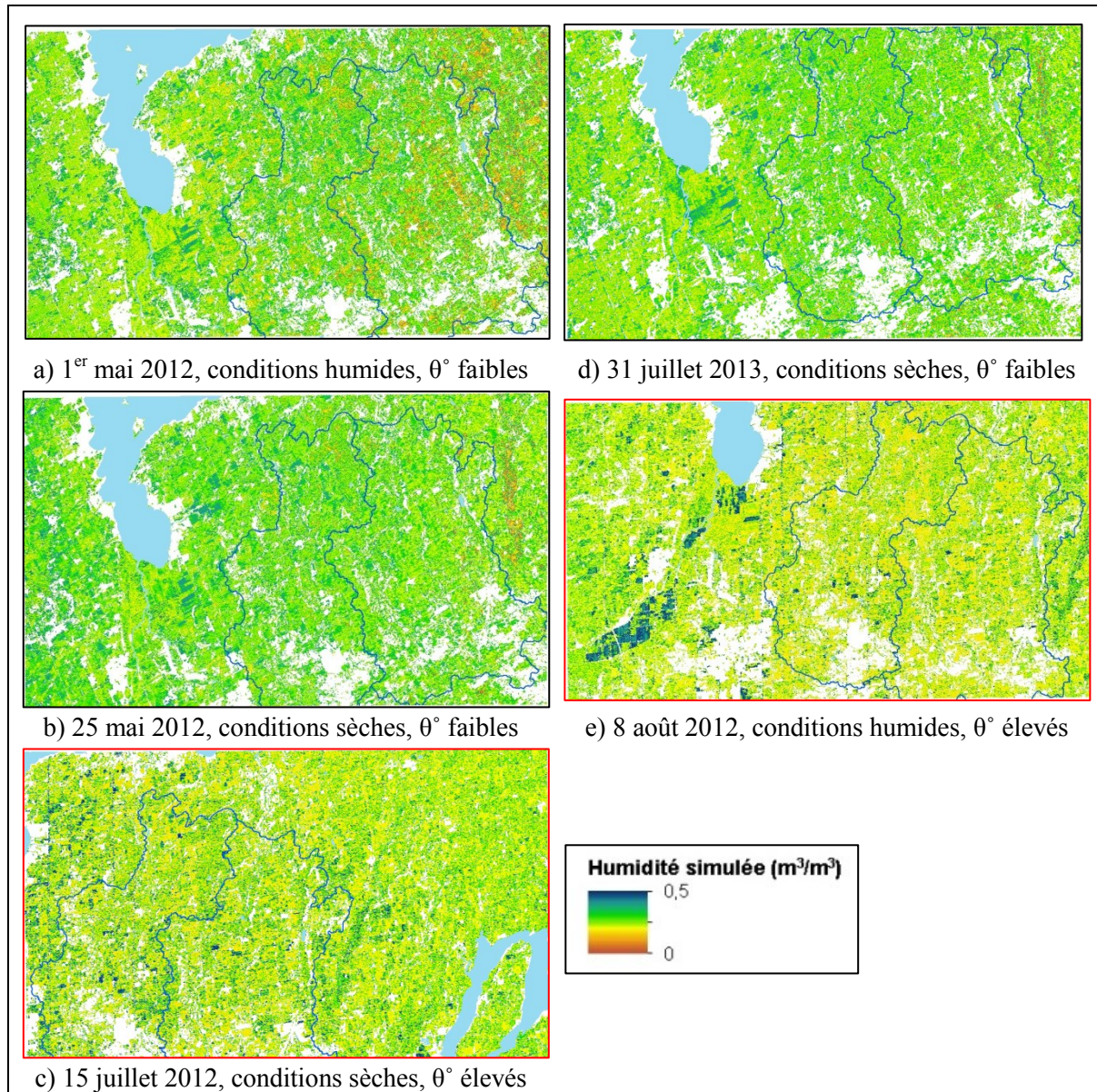


Figure 5.33 Carte d'estimation de l'humidité du sol avec $\text{RNA}_{\theta^\circ, \sigma^\circ \text{Surf_HH}}$.

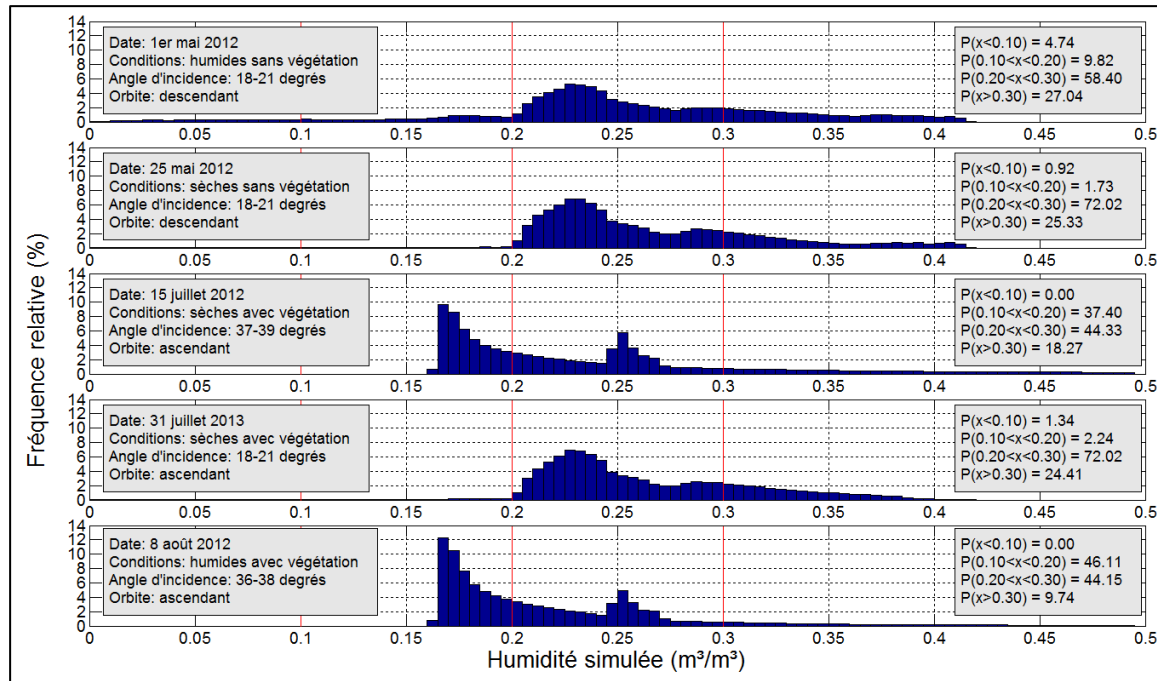


Figure 5.34 Histogrammes de l'humidité estimée avec $RNA_{\theta^{\circ}, \sigma^{\circ} \text{Surf_HH}}$ pour les sous-bassins versants Baldwin et Pepperlaw

Les résultats contradictoires obtenus avec l'angle d'incidence et la rétrodiffusion de surface en polarisation HH comme intrants aux RNA sont attribuables à plusieurs facteurs. En ne conservant que la contribution de la surface au signal rétrodiffusé, l'intensité de la rétrodiffusion de surface est fonction de la quantité de végétation présente. En effet, à mesure que la végétation se développe la dépolarisation du signal et la diffusion de volume qu'elle entraîne augmentent, ce qui cause une diminution de $\sigma^{\circ}_{\text{Surf_HH}}$. Ainsi, les valeurs de rétrodiffusion de surface ainsi que l'humidité estimée sont généralement plus élevées en absence de végétation abondante, soit le 1^{er} et le 25 mai 2012. Par contre, l'humidité élevée combinée à la forte rugosité de surface associée à des sols agricoles nus a généré une rétrodiffusion de surface supérieure aux valeurs contenues dans le jeu de données utilisé pour développer les RNA. Les RNA n'ayant pas appris ces cas de figure, l'humidité du sol estimée à partir de ces données est d'environ $0,10 \text{ m}^3/\text{m}^3$, ce qui est irréaliste. Dans le cas du 31 juillet 2013, l'humidité estimée est comparable à celle du 25 mai 2012 et semble indiquer que les RNA ont la capacité d'estimer l'humidité du sol malgré une végétation abondante. Une végétation abondante aurait toutefois entraîné une rétrodiffusion de surface nettement plus faible que celle mesurée le 25 mai 2012, ce qui n'est pas le cas. En effet, la rétrodiffusion

mesurée est du même ordre de grandeur pour les deux dates. Cette forte rétrodiffusion de surface peut être due au fait qu'une récolte a possiblement eu lieu dans les jours précédents. Dans ce cas, la quantité de végétation présente serait similaire. D'autre part, 51 mm de pluie ont été enregistrés ce jour-là. On peut donc émettre l'hypothèse qu'il pleuvait et que la végétation était mouillée lors de l'acquisition, ce qui a pu être interprété à tort comme de la rétrodiffusion de surface lors de la décomposition polarimétrique.

L'obtention de cartes plus sèches obtenues pour les images du 15 juillet et du 8 août 2012 est le résultat de la combinaison d'une végétation abondante et d'angles d'incidence élevés. En effet, la quantité de rétrodiffusion provenant de la surface est réduite de manière importante de par la diffusion de volume produite par la végétation abondante, ce qui est amplifié par les angles d'incidence élevés. Ainsi, des images $\sigma^{\circ}_{\text{Surf_HH}}$ de faible intensité et des cartes d'humidité estimée généralement plus sèches ont été obtenues pour ces dates.

Lorsque comparées aux cartes obtenues avec les $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}\text{HH}}$, les différences sont principalement fonction des angles d'incidence. En effet, l'humidité estimée est généralement plus élevée à angles d'incidence faibles alors qu'elle est moindre à angles d'incidence élevés. Les différences sont les plus marquées entre les cartes du 25 mai 2012. En effet, l'humidité estimée pour les sols mal drainés sont les plus humides sur la carte obtenue avec les $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}\text{HH}}$ alors qu'avec les $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}\text{Surf_HH}}$, l'humidité estimée est relativement homogène sur l'ensemble du territoire. Bien qu'aucune observation n'ait été faite sur le terrain, la cartographie de l'utilisation du sol (figure 4.9) ainsi que les photographies aériennes révèlent que ces sols sont principalement couverts de forêt et de milieux humides. Ainsi, on peut affirmer que ces zones sont davantage humides et que les résultats obtenus avec les $\text{RNA}_{\theta^{\circ}, \sigma^{\circ}\text{HH}}$ sont plus fiables dans la mesure où l'humidité du sol estimée présente une variabilité spatiale corrélée à la qualité du drainage des sols. Les valeurs de m_{ve} plus élevées sont par contre davantage liées à la rétrodiffusion plus importante provenant de la végétation que du sol.

La distribution spatiale de l'incertitude sur les estimations suit la même tendance que pour les autres combinaisons d'intrants : elle est généralement homogène et sa magnitude est corrélée aux valeurs de rétrodiffusion. En effet, l'incertitude est nettement plus élevée pour les valeurs extrêmes de $\sigma^{\circ}_{\text{Surf_HH}}$ qui ne sont pas comprises dans le jeu de données d'entraînement. Les

histogrammes de fréquence de la figure 5.35 indiquent que l'incertitude pour les sous-bassins versants Baldwin et Pepperlaw se concentre en-deçà de $0,05 \text{ m}^3/\text{m}^3$ pour l'ensemble des cartes.

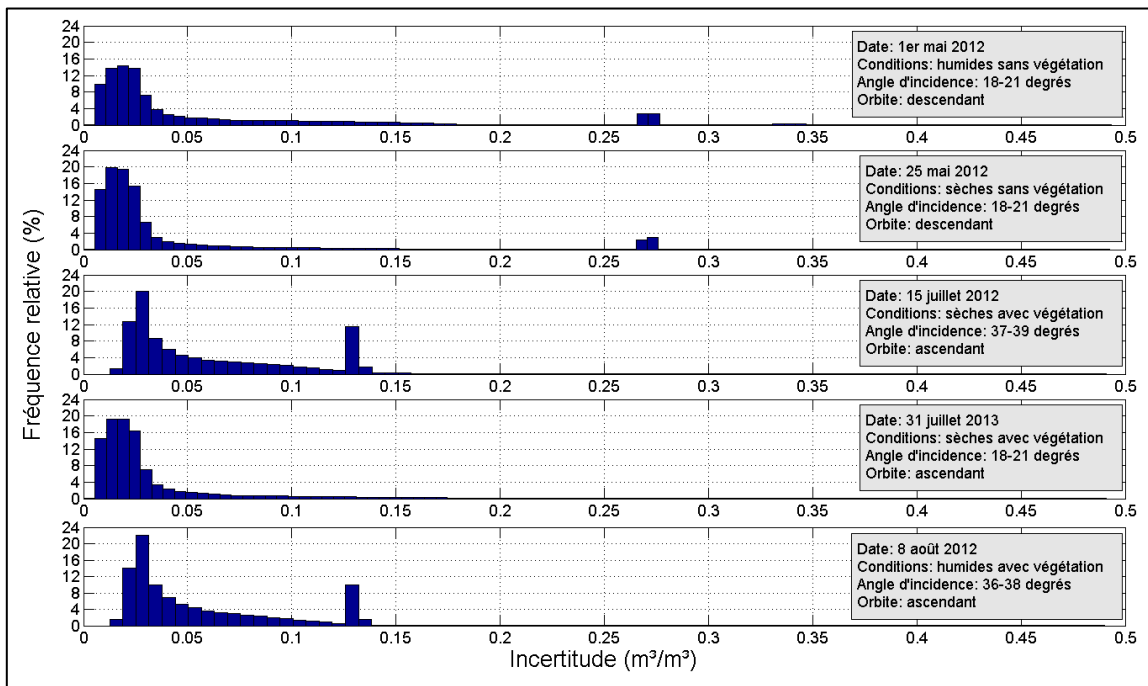


Figure 5.35 Histogrammes d'incertitude pour les sous-bassins versants Baldwin et Pepperlaw avec $\text{RNA}_{\theta^\circ, \sigma^\circ \text{Surf_HH}}$

Bien que l'incertitude des $\text{RNA}_{\theta^\circ, \sigma^\circ \text{Surf_HH}}$ soit moindre, l'utilisation de $\sigma^\circ \text{Surf_HH}$ en remplacement de $\sigma^\circ \text{HH}$ comme intrant ne permet pas d'obtenir de meilleurs résultats. En effet, le gain réalisé en termes d'incertitude ne compense pas la diminution des performances obtenues, notamment au niveau du coefficient de Pearson et de l'erreur relative moyenne.

Au final, c'est l'approche *Polarisation simple et double* et plus précisément les RNA ayant pour intrants l'angle d'incidence ainsi que les coefficients de rétrodiffusion en polarisation HH et HV qui permettent d'obtenir les meilleurs résultats parmi l'ensemble des combinaisons analysées.

5.3 Influence de la taille de l'échantillon

Dans le but d'évaluer l'influence de la taille de l'échantillon sur les performances du meilleur RNA pouvant être obtenu ainsi que sur les cartes d'estimation de l'humidité et d'incertitude produites, les statistiques extraites des images acquises en mode Standard et Large ont été ajoutées à celles provenant des images polarimétriques. En ajoutant les échantillons provenant

des images acquises en mode Standard et Large en polarisation HH et HV, la taille de l'échantillon est passée de 119 à 208 puis à 534 avec celles en polarisation HH uniquement.

Les régressions linéaires (figure 5.36) ainsi que le sommaire des performances (tableau 5.16) mettent en évidence plusieurs éléments. L'utilisation de l'échantillon de 208 données n'a pas affecté de manière significative ($p\text{-value} > 0,05$ pour les coefficients de Pearson) les performances ni le comportement général du modèle.

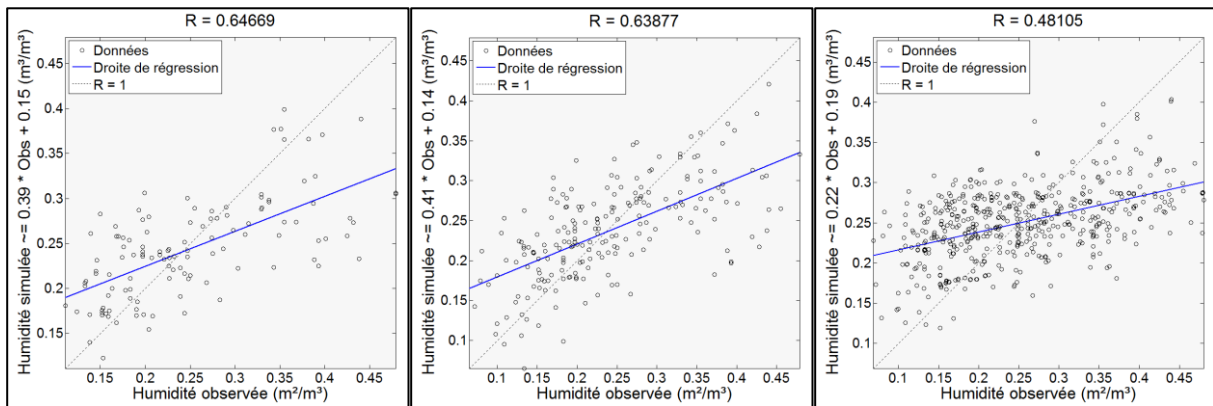


Figure 5.36 Régressions linéaires des meilleurs $\text{RNA}_{\theta^{\circ}, \sigma^{\circ} \text{HH}}$ utilisant 119 (gauche), 208 (centre) et 534 (droite) données

Tableau 5.16 Performances du meilleur RNA en fonction de la taille de l'échantillon

Taille de l'échantillon	Coefficient de Pearson (r)	RMSE (m^3/m^3)	MAE (m^3/m^3)	MRE (%)
119	0,647	0,069	0,053	22,183
208	0,639	0,071	0,054	25,455
534	0,481	0,082	0,068	31,876

En effet, les valeurs observées en deçà de $0,30 \text{ m}^3/\text{m}^3$ sont autant sous-estimées que surestimées alors qu'au-delà, elles sont systématiquement sous-estimées. Le changement le plus notable est l'augmentation de 3,5 % de l'erreur relative moyenne. Par contre, le domaine de valeurs d'humidité observée de l'échantillon de taille 208 est de $[0,069 : 0,480]$ comparativement à $[0,111 : 0,479]$ pour celui de 119. Ce qui signifie que le modèle a la capacité de générer des valeurs minimales plus faibles et ainsi capter une gamme d'humidité plus vaste.

L'utilisation de l'échantillon de 534 données a par contre dégradé de manière significative ($p\text{-value} < 0,05$ pour les coefficients de Pearson) les performances comparativement à

l'utilisation de 119 et 208 données. En effet, le coefficient de Pearson est de 0,48 alors que l'erreur relative moyenne est de 31,9 %, en plus de la pente de la droite de régression qui a diminué de manière importante. Plusieurs facteurs peuvent avoir contribué à cette dégradation des performances. En effet, les données qui ont été ajoutées au jeu de données de 208 couples proviennent d'images acquises majoritairement à des angles d'incidence élevés (figure 5.37).

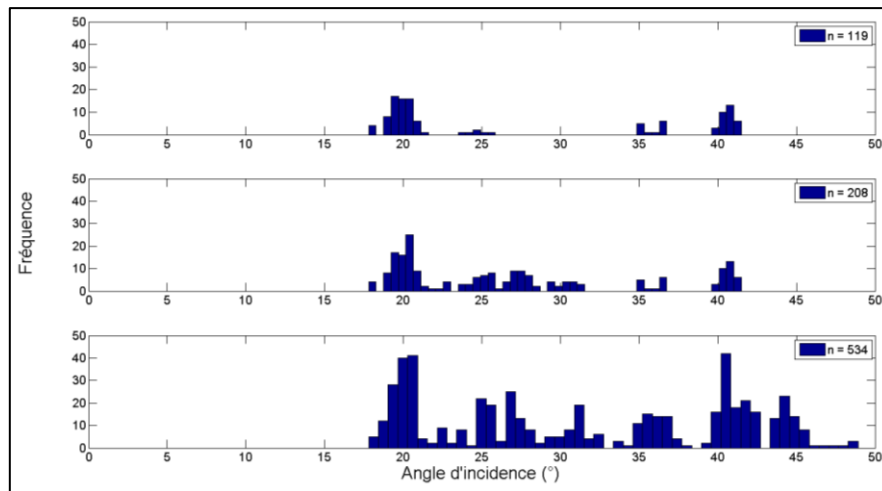


Figure 5.37 Histogrammes de distribution des valeurs d'angles d'incidence

5.3.1 Cartes d'estimation de l'humidité du sol

Les statistiques des 30 meilleurs $RNA_{\theta^*, \sigma^{*HH}}$ obtenus avec les différents jeux de données sont présentées au tableau 5.17. Aucune différence significative ($p\text{-values} > 0,05$) n'est observée entre les différents coefficients de Pearson. Le plus faible demeure celui obtenu avec 534 données alors que l'écart entre ceux obtenus avec 119 et 208 données est moindre comparativement aux résultats du meilleur RNA présentés au tableau 5.16.

Tableau 5.17 Sommaire des statistiques des 30 meilleurs RNA en fonction de la taille de l'échantillon

Taille de l'échantillon	Moyenne				Domaine de valeur							
	Coefficient de Pearson (r)	RMSE (m³/m³)	MAE (m³/m³)	MRE (%)	Coefficient de Pearson (r)		RMSE (m³/m³)		MAE (m³/m³)		MRE (%)	
					Min	Max	Min	Max	Min	Max	Min	Max
119	0,592	0,073	0,057	24,7	0,55	0,66	0,068	0,076	0,051	0,061	22,0	26,7
208	0,559	0,077	0,061	28,6	0,53	0,64	0,071	0,079	0,054	0,063	25,5	30,0
534	0,468	0,083	0,069	32,1	0,44	0,50	0,082	0,084	0,067	0,070	31,3	33,1

En appliquant les 30 meilleurs RNA, les cartes d'humidité estimée présentées à la figure 5.38 ont été obtenues. Les patrons de distribution de l'humidité sont restés les mêmes indépendamment de la taille de l'échantillon. La variation des valeurs d'humidité du sol entre les différentes tailles des échantillons est principalement liée aux angles d'incidence. En effet, avec une augmentation de la taille de l'échantillon, l'humidité estimée augmente pour les images acquises à faibles angles d'incidence alors qu'elle diminue pour celles acquises à angles d'incidence élevés.

En conclusion, parmi les RNA utilisant l'angle d'incidence et σ_{HH}° comme seuls intrants, les résultats obtenus indiquent que l'utilisation de l'échantillon de 208 données est celui offrant le meilleur compromis entre performance et robustesse. En effet, les gains réalisés au niveau de l'étendue des valeurs d'humidité estimées ainsi que les écarts-type moindres compensent largement le coefficient de Pearson inférieur de 0,04 ainsi que l'erreur relative moyenne supérieure de 3,3 %.

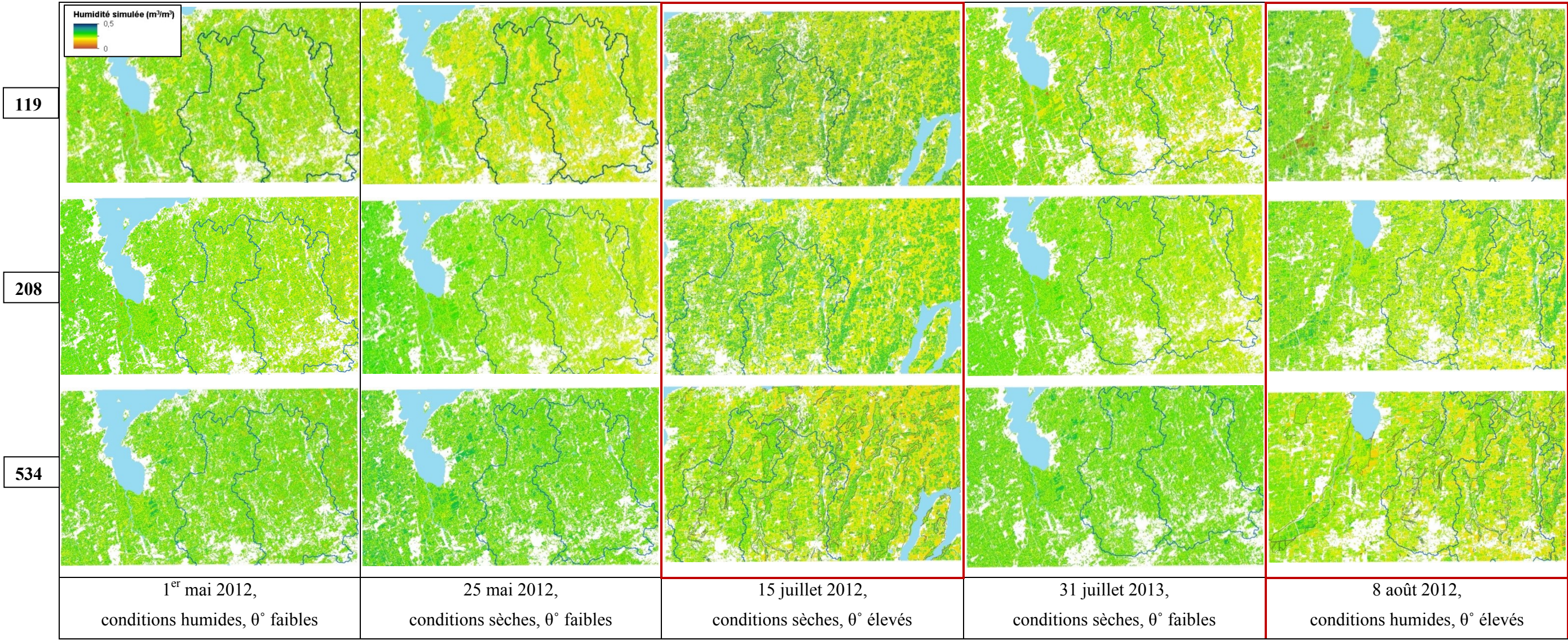


Figure 5.38 Carte d'estimation de l'humidité du sol RNA θ° , σ° HH en fonction de la taille de l'échantillon

CHAPITRE 6 CONCLUSION

Le projet de recherche était basé sur l'utilisation de données RADARSAT-2 et de réseaux de neurones artificiels de type perceptrons multicouches afin d'estimer l'humidité du sol en surface pour des sols agricoles nus ou faiblement végétalisés. Il visait à vérifier si l'utilisation de paramètres polarimétriques comme intrants aux RNA apporte une augmentation significative de la précision et une réduction de l'erreur sur l'estimation de l'humidité du sol. Afin de vérifier cette hypothèse, les approches *Polarisation simple et double* et *Polarimétrique* ont été évaluées. En plus des données radar et des paramètres polarimétriques, de l'information auxiliaire a aussi été utilisée comme intrant aux RNA. Plus spécifiquement, les objectifs étaient de déterminer l'approche la plus avantageuse dans un contexte opérationnel, de quantifier l'incertitude sur les estimations d'humidité du sol et de produire des cartes d'humidité du sol et d'incertitude de l'estimation.

Les résultats obtenus ont permis d'atteindre l'ensemble des objectifs et de tirer les conclusions suivantes :

- Parmi les différentes combinaisons analysées utilisant les données en polarisation simple et double, les meilleurs résultats ont été obtenus avec l'angle d'incidence (θ°) et les coefficients de rétrodiffusion radar en polarisation HH (σ_{HH}°) et HV (σ_{HV}°) comme intrants.
- L'incertitude sur l'estimation de l'humidité du sol est étroitement liée aux données d'entraînement. En effet, on a pu observer une incertitude élevée lorsque de l'extrapolation a eu lieu, et ce, pour l'ensemble des combinaisons d'intrants analysées. De nombreux auteurs ont aussi observé ce comportement (ex. [Baghdadi *et al.*, 2012b]).
- L'utilisation de variables statiques dans le temps comme intrants affecte négativement et de façon importante les cartes d'estimation obtenues malgré des performances élevées (statistiques). En effet, l'utilisation de données auxiliaires telles que la fraction de sable et d'argile a permis d'obtenir le meilleur coefficient de Pearson (0,723) pour l'approche de polarisation simple et double. L'analyse visuelle des cartes d'humidité estimée et des statistiques associées a toutefois démontré une insensibilité des RNA à l'humidité du sol. Ce constat met en évidence que l'analyse de critères de performances n'est pas suffisante

pour l'évaluation de la qualité d'un RNA. En effet, la production de cartes permet d'évaluer visuellement la cohérence de la distribution spatiale de l'humidité du sol estimée et d'identifier les zones où l'incertitude est particulièrement élevée.

- Parmi les combinaisons utilisant les paramètres polarimétriques, le RNA utilisant l'angle d'incidence ainsi que la puissance de diffusion de surface et de volume comme intrants est celui pour lequel les meilleures performances ont été obtenues.
- L'utilisation de données polarimétriques permet une amélioration de la précision ainsi qu'une réduction de l'erreur sur les estimations d'humidité du sol. Toutefois, ces avantages par rapport à l'approche *Polarisation simple et double* ne sont pas statistiquement significatifs. Ce résultat est en accord avec les études de Baghdadi *et al.* [2012b; 2016].
- Compte tenu que ce projet a été réalisé dans un contexte de développement d'une approche opérationnelle et du coût supplémentaire associé aux images polarimétriques, l'approche *Polarisation simple et double* est considérée comme la plus avantageuse.

La réalisation des cartes d'incertitude a permis de mettre en évidence l'importance d'utiliser un jeu de données d'entraînement représentatif des conditions où les RNA seront appliqués. En effet, bien que certains patrons de distribution présents sur les cartes semblent indiquer des anomalies, les cartes d'incertitude permettent de la quantifier. Cette information peut donc être utilisée afin d'identifier les conditions les plus problématiques et ainsi être en mesure d'ajouter des données d'entraînement, ce qui à terme permettrait d'améliorer les performances.

Ce travail a permis de mettre en évidence le potentiel, mais aussi les difficultés rencontrées dans la mise en œuvre d'un modèle de rétrodiffusion basé sur l'approche des RNA. Il est par conséquent suggéré de poursuivre les recherches dans cette direction, notamment en s'inspirant des recommandations suivantes :

- De nombreuses variantes existent en termes d'algorithme d'apprentissage, d'optimisation d'architecture et d'intrants, etc. Une approche de plus en plus courante utilise la modélisation à bases physiques pour générer des observations synthétiques et combler les trous laissés par le manque d'observations réelles disponibles. Cette méthode permet de générer un grand nombre de données d'«observations» (*ex* : plus de 10 000) couvrant l'ensemble du domaine des conditions potentiellement rencontrées sur le terrain (*ex* : humidité du sol, végétation, angles d'incidence, etc.) pour l'entraînement des RNA.

Toutefois, les observations *in situ* actuellement disponibles dans cette étude ne sont pas en nombre suffisamment important pour couvrir toutes les conditions, de telle sorte qu'une augmentation de la taille de l'échantillon a amené une diminution de la performance des RNA. Toutefois, l'objectif de ce projet n'était pas de développer le meilleur des RNA mais plutôt d'évaluer l'influence sur les performances et sur l'incertitude de différentes variables.

- Toujours dans un contexte d'amélioration des RNA pour l'estimation de l'humidité du sol, il est suggéré d'établir des classes d'humidité pour l'entraînement des RNA, plutôt que d'utiliser directement les observations, afin de diminuer la variabilité et de prendre en considération les difficultés associées à l'estimation de l'humidité du sol en conditions très sèches et très humides. La capacité de cette approche d'augmenter les performances des RNA est toutefois conditionnelle à ce qu'un échantillon de grande taille soit disponible afin que chaque classe d'humidité du sol contienne un nombre d'échantillons suffisamment élevé.
- Au lieu d'utiliser directement la variable $\sigma^{\circ}_{\text{Surf_HH}}$ dans le RNA, il y aurait lieu de la corriger pour prendre en compte que la végétation masque partiellement la surface et entraîner le RNA avec ces valeurs modifiées.
- On pourrait aussi inclure une estimation de la densité de végétation dans l'apprentissage du RNA, pour autant qu'on soit capable d'estimer cette valeur. L'utilisation de P_v , qui est sensible à la présence de végétation, pourrait s'avérer une avenue possible, de même que de recourir à de l'imagerie à haute résolution dans le domaine du visible (ex : MODIS, Sentinel-2).

LISTE DES RÉFÉRENCES

- Ali, S., Ghosh, N. C. et Singh, R. (2010) Rainfall–runoff simulation using a normalized antecedent precipitation index. *Hydrological Sciences Journal*, vol. 55, n°2, p. 266–274.
- Altese, E., Bolognani, O., Mancini, M. et Troch, P. A. (1996) Retrieving Soil Moisture Over Bare Soil from ERS 1 Synthetic Aperture Radar Data: Sensitivity Analysis Based on a Theoretical Surface Scattering Model and Field Data. *Water Resources Research*, vol. 32, n°3, p. 653–661.
- Amari, S., Murata, N., Müller, K.-R., Finke, M. et Yang, H. H. (1997) Asymptotic statistical theory of overtraining and cross-validation. *Neural Networks, IEEE Transactions on*, vol. 8, n°5, p. 985–996.
- ASCE Task Committee (2000) Artificial Neural Networks in Hydrology. I: Preliminary Concepts. *Journal of Hydrologic Engineering*, vol. 5, n°2, p. 115–123.
- Attema, E. P. W. et Ulaby, F. T. (1978) Vegetation modeled as a water cloud. *Radio science*, vol. 13, n°2, p. 357–364.
- Baghdadi, N., Cerdan, O., Zribi, M., Auzet, V., Darboux, F., El Hajj, M. et Kheir, R. B. (2008) Operational performance of current synthetic aperture radar sensors in mapping soil surface characteristics in agricultural environments: application to hydrological and erosion modelling. *Hydrological Processes*, vol. 22, n°1, p. 9–20.
- Baghdadi, N., Cresson, R., El Hajj, M., Ludwig, R. et La Jeunesse, I. (2012) Estimation of soil parameters over bare agriculture areas from C-band polarimetric SAR data using neural networks. *Hydrology and Earth System Sciences*, vol. 16, n°6, p. 1607–1621.
- Baghdadi, Nicolas, Cresson, R., Pottier, E., Aubert, M., Mehrez, M., Jacome, A. et Benabdallah, S. (2012) A Potential Use for the C-Band Polarimetric SAR Parameters to Characterize the Soil Surface Over Bare Agriculture Fields. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, n°10, p. 3844–3858.
- Baghdadi, N., Gaultier, S. et King, C. (2002) Retrieving surface roughness and soil moisture from synthetic aperture radar (SAR) data using neural networks. *Canadian Journal of Remote Sensing*, vol. 28, n°5, p. 701–711.
- Baghdadi, N., Gherboudj, I., Zribi, M., Sahebi, M., King, C. et Bonn, F. (2004) Semi-empirical calibration of the IEM backscattering model using radar images and moisture and roughness field measurements. *International Journal of Remote Sensing*, vol. 25, n°18, p. 3593–3623.
- Baghdadi, N., Holah, N. et Zribi, M. (2006a) Calibration of the Integral Equation Model for SAR data in C-band and HH and VV polarizations. *International Journal of Remote Sensing*, vol. 27, n°4, p. 805–816.
- Baghdadi, N., Holah, N. et Zribi, M. (2006b) Soil moisture estimation using multi-incidence and multi-polarization ASAR data. *International Journal of Remote Sensing*, vol. 27, n°10, p. 1907–1920.

- Baghdadi, N., King, C., Bourguignon, A. et Remond, A. (2002) Potential of ERS and Radarsat data for surface roughness monitoring over bare agricultural fields: Application to catchments in Northern France. *International Journal of Remote Sensing*, vol. 23, n°17, p. 3427-3442.
- Baghdadi, N., King, C., Chanzy, A. et Wigneron, J. P. (2002) An empirical calibration of the integral equation model based on SAR data, soil moisture and surface roughness measurement over bare soils. *International Journal of Remote Sensing*, vol. 23, n°20, p. 4325-4340.
- Baghdadi, N. N., El Hajj, M., Zribi, M. et Fayad, I. (2016) Coupling SAR C-Band and Optical Data for Soil Moisture and Leaf Area Index Retrieval Over Irrigated Grasslands. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, n°3, p. 1229-1243.
- Baghdadi, N., Paillou, P., Grandjean, G., Dubois, P. et Davidson, M. (2000) Relationship between profile length and roughness variables for natural surfaces. *International Journal of Remote Sensing*, vol. 21, n°17, p. 3375-3381.
- Baldwin, D. J., Desloges, J. R. et Band, L. E. (2011) Chapter 2: Physical Geography of Ontario. *Ecology of a managed terrestrial landscape: patterns and processes of forest landscapes in Ontario* (p. 12-19). Vancouver, Canada : University of British Columbia Press.
- Barrett, B., Dwyer, E. et Whelan, P. (2012) Sensitivity of spaceborne radar to near-surface soil moisture in grasslands across southern Ireland. *Irish Geography*, vol. 45, n°2, p. 131-149.
- Beck, H. E., de Jeu, R. A. M., Schellekens, J., van Dijk, A. I. J. M. et Bruijnzeel, L. A. (2009) Improving Curve Number Based Storm Runoff Estimates Using Soil Moisture Proxies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, n°4, p. 250-259.
- Blumberg, D. G., Freilikher, V., Lyalko, I. V., Vulfson, L. D., Kotlyar, A. L., Shevchenko, V. N. et Ryabokononko, A. D. (2000) Soil moisture (water-content) assessment by an airborne scatterometer: The Chernobyl disaster area and the Negev Desert. *Remote Sensing of Environment*, vol. 71, n°3, p. 309-319.
- Boisvert, J. B., Gwyn, Q. H. J., Brisco, B., Major, D. J. et Brown, R. J. (1995) Evaluation of Soil Moisture Estimation Techniques and Microwave Penetration Depth for Radar Applications. *Canadian Journal of Remote Sensing*, vol. 21, n°2, p. 110-123.
- Bourgeau-Chavez, L. L., Leblon, B., Charbonneau, F. et Buckley, J. R. (2013) Evaluation of polarimetric Radarsat-2 SAR data for development of soil moisture retrieval algorithms over a chronosequence of black spruce boreal forests. *Remote Sensing of Environment*, vol. 132, p. 71-85.
- Bowden, G. J., Maier, H. R. et Dandy, G. C. (2002) Optimal division of data for neural network models in water resources applications: DIVISION OF DATA FOR NEURAL NETWORK MODELS. *Water Resources Research*, vol. 38, n°2, p. 2-1-2-11.

- Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z. et Hasenauer, S. (2010) Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrology and Earth System Sciences*, vol. 14, n°10, p. 1881-1893.
- Bruckler, L., Witono, H. et Stengel, P. (1988) Near surface soil moisture estimation from microwave measurements. *Remote Sensing of Environment*, vol. 26, n°2, p. 101-121.
- Bryant, R., Moran, M. S., Thoma, D. P., Holifield Collins, C. D., Skirvin, S., Rahman, M., Slocum, K., Starks, P., Bosch, D. et Gonzalez Dugo, M. P. (2007) Measuring Surface Roughness Height to Parameterize Radar Backscatter Models for Retrieval of Surface Soil Moisture. *IEEE Geoscience and Remote Sensing Letters*, vol. 4, n°1, p. 137-141.
- Capehart, W. J. et Carlson, T. N. (1997) Decoupling of surface and near-surface soil water content: A remote sensing perspective. *Water Resources Research*, vol. 33, n°6, p. 1383-1395.
- Cassardo, C., Balsamo, G. P., Cacciamani, C., Cesari, D., Paccagnella, T. et Pelosini, R. (2002) Impact of soil surface moisture initialization on rainfall in a limited area model: a case study of the 1995 South Ticino flash flood. *Hydrological Processes*, vol. 16, n°6, p. 1301-1317.
- Centre canadien de télédétection (1997) Ressources didactiques de télédétection radar.
- Charbonneau, F., Trudel, M. et Fernandes, R. (2005) Use of Dual-Polarization and Multi-Incidence SAR for soil permeability mapping, *Atelier Advanced Synthetic Aperture Radar (ASAR) 2005*, Saint-Hubert, Canada, novembre 2005
- Chauhan, N. S., Lang, R. H. et Ranson, K. J. (1991) Radar modeling of a boreal forest. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, n°4, p. 627-638.
- Chow, V. T., Maidment, D. R. et Mays, L. W. (1988) *Applied hydrology*. McGraw-Hill series in water resources and environmental engineering. New York : McGraw-Hill.
- Cloude, S. R. et Pottier, E. (1997) An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, n°1, p. 68-78.
- Cui, Y., Yamaguchi, Y., Yang, J. et Kobayashi, H. (2012) On exact model-based scattering decomposition of polarimetric SAR data. *2012 International Symposium on Antennas and Propagation (ISAP)*, p. 106-109.
- Dawson, M. S., Fung, A. K. et Manry, M. T. (1997) A robust statistical-based estimator for soil moisture retrieval from radar measurements. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, n°1, p. 57-67.
- Decagon Devices, Inc. (2001) *EC-5 Soil Moisture Sensor User's Manual*, Version 2.
- Del Frate, F., Ferrazzoli, P. et Schiavon, G. (2003) Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks. *Remote sensing of environment*, vol. 84, n°2, p. 174-183.
- Dobson, M. C. et Ulaby, F. (1981) Microwave backscatter dependence on surface roughness, soil moisture, and soil texture: Part III-Soil tension. *Geoscience and Remote Sensing, IEEE Transactions on*, n°1, p. 51-61.

- Dobson, M. C. et Ulaby, F. T. (1986) Active microwave soil moisture research. *Geoscience and Remote Sensing, IEEE Transactions on*, n°1, p. 23–36.
- Dobson, M. C., Ulaby, F. T., Hallikainen, M. T. et El-rayes, M. A. (1985) Microwave Dielectric Behavior of Wet Soil-Part II: Dielectric Mixing Models. *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-23, n°1, p. 35–46.
- Dubois, P. C., van Zyl, J. et Engman, T. (1995a) Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, n°4, p. 915–926.
- Dubois, P. C., van Zyl, J. et Engman, T. (1995b) Corrections to « Measuring Soil Moisture with Imaging Radars ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, n°6, p. 1340.
- Ecoplans Limited (2007) A Study of the Past, Present and Future of Water Management on the Trent- Severn Waterway National Historic Site of Canada. Water Management Program. (p. 33).
- ESA (2011) PolSarPro. ESA.
- ESA (2012) NEST Release 4C-1.1, Next ESA SAR Toolbox.
- ESRI (2011) ArcGIS Desktop 10.1. Redlands, CA : Environmental Systems Research Institute.
- Freeman, A. et Durden, S. L. (1998) A three-component scattering model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, n°3, p. 963–973.
- Fung, A. K. (1994) Microwave scattering and emission models and their applications. The Artech House remote sensing library. Boston, États-Unis : Artech House.
- Fung, A. K., Li, Z. et Chen, K.-S. (1992) Backscattering from a randomly rough dielectric surface. *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 30, n°2, p. 356–369.
- Gouvernement du Canada (2007) Données numériques d'élévation du Canada, Niveau 1 Spécifications de produit Édition 3.0. Ressources naturelles Canada Centre d'information topographique.
- Gouvernement du Canada (2016) http://climat.meteo.gc.ca/historical_data/search_historic_data_f.html. Données historiques.
- Grabs, T., Seibert, J., Bishop, K. et Laudon, H. (2009) Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. *Journal of Hydrology*, vol. 373, n°1–2, p. 15–23.
- Hajnsek, I., Cloude, S. R., Lee, J.-S. et Pottier, E. (2000) Inversion of surface parameters from polarimetric SAR data, vol. 3, p. 1095–1097. IEEE.
- Hajnsek, I., Jagdhuber, T., Schon, H. et Papathanassiou, K. P. (2009) Potential of Estimating Soil Moisture Under Vegetation Cover by Means of PolSAR. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, n°2, p. 442–454.

- Hajnsek, I., Pottier, E. et Cloude, S. R. (2003) Inversion of surface parameters from polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, n°4, p. 727-744.
- Hallikainen, M. T., Ulaby, F. T., Dobson, M. C., El-Rayes, M. A. et Wu, L.-K. (1985) Microwave dielectric behavior of wet soil-part 1: empirical models and experimental observations. *IEEE Transactions on Geoscience and Remote Sensing*, n°1, p. 25-34.
- Harrington, P. de B. (1993) Sigmoid transfer functions in backpropagation neural networks. *Analytical Chemistry*, vol. 65, n°15, p. 2167-2168.
- Haykin, S. (1994) *Neural networks: a comprehensive foundation*. New York, États-Unis : Macmillan College Publishing Company Inc.
- Heggen, R. J. (2001) Normalized antecedent precipitation index. *Journal of hydrologic Engineering*, vol. 6, n°5, p. 377-381.
- Henderson, F. M., Lewis, A. J. et American Society for Photogrammetry and Remote Sensing (dir.) (1998) *Principles and applications of imaging radar. Manual of remote sensing*, third edition. New York : J. Wiley.
- Holah, N., Baghdadi, N., Zribi, M., Bruand, A. et King, C. (2005) Potential of ASAR/ENVISAT for the characterization of soil surface parameters over bare agricultural fields. *Remote Sensing of Environment*, vol. 96, n°1, p. 78-86.
- Karam, M. A., Amar, F., Fung, A. K., Mougin, E., Lopes, A., Le Vine, D. M. et Beaudoin, A. (1995) A microwave polarimetric scattering model for forest canopies based on vector radiative transfer theory. *Remote Sensing of Environment*, vol. 53, n°1, p. 16-30.
- Karam, M. A., Fung, A. K., Lang, R. H. et Chauhan, N. S. (1992) A microwave scattering model for layered vegetation. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, n°4, p. 767-784.
- Kim, Y. et van Zyl, J. (2004) Vegetation effects on soil moisture estimation. *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, p. 800-802.
- Kim, Y. et van Zyl, J. J. (2009) A Time-Series Approach to Estimate Soil Moisture Using Polarimetric Radar Data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, n°8, p. 2519-2527.
- Kong, X. et Dorling, S. R. (2008) Near-surface soil moisture retrieval from ASAR Wide Swath imagery using a Principal Component Analysis. *International Journal of Remote Sensing*, vol. 29, n°10, p. 2925-2942.
- Kornelsen, K. C. et Coulibaly, P. (2013) Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. *Journal of Hydrology*, vol. 476, p. 460-489.
- Kseneman, M., Gleich, D. et Potočník, B. (2012) Soil-moisture estimation from TerraSAR-X data using neural networks. *Machine Vision and Applications*, vol. 23, n°5, p. 937-952.

- Lakhankar, T., Ghedira, H. et Khanbilvardi, R. (2006) Spatial Mapping of Soil Moisture Using RADARSAT-1 Data. Proceedings of the ASPRS 2006 Annual Conference, Reno, NV.
- Le groupe de travail sur les Pédo-paysages du Canada, (2010) Pédo-paysage du Canada version 3.2. Agriculture et Agroalimentaire Canada. (carte numérique et base de données à l'échelle de 1 / 1 million).
- Lee, J.-S. et Pottier, E. (2009) Polarimetric radar imaging: from basics to applications. Optical science and engineering. Boca Raton : CRC Press.
- Levenberg, K. (1944) A method for the solution of certain non-linear problems in least squares. Quarterly of applied mathematics, vol. 2, p. 164-168.
- Lievens, H., Vernieuwe, H., Álvarez-Mozos, J., De Baets, B. et Verhoest, N. E. C. (2009) Error in Radar-Derived Soil Moisture due to Roughness Parameterization: An Analysis Based on Synthetical Surface Profiles. Sensors, vol. 9, n°2, p. 1067-1093.
- Looney, C. G. (1996) Advances in feedforward neural networks: demystifying knowledge acquiring black boxes. IEEE Transactions on Knowledge and Data Engineering, vol. 8, n°2, p. 211-226.
- Lopes, A., Touzi, R. et Nezry, E. (1990) Adaptive speckle filters and scene heterogeneity. IEEE Transactions on geoscience and remote sensing, vol. 28, n°6, p. 992-1000.
- MacDonald, Dettwiler and Associates Ltd. (2014) RADARSAT-2 PRODUCT DESCRIPTION.
- Maier, H. R. et Dandy, G. C. (1998) The effect of internal parameters and geometry on the performance of back-propagation neural networks: an empirical study. Environmental Modelling & Software, vol. 13, n°2, p. 193-209.
- Maier, H. R. et Dandy, G. C. (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental modelling & software, vol. 15, n°1, p. 101-124.
- Marquardt, D. W. (1963) An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics, vol. 11, n°2, p. 431-441.
- Maurer, E. P., Lettenmaier, D. P. et Mantua, N. J. (2004) Variability and potential sources of predictability of North American runoff. Water Resources Research, vol. 40, n°9.
- McNairn, H. et Brisco, B. (2004) The application of C-band polarimetric SAR for agriculture: a review. Canadian Journal of Remote Sensing, vol. 30, n°3, p. 525-542.
- McNairn, H., Merzouki, A. et Pacheco, A. (2010) Estimating surface soil moisture using Radarsat-2. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, vol. 38, n°8, p. 576-579.
- Merz, B. et Plate, E. J. (1997) An analysis of the effects of spatial variability of soil and soil moisture on runoff. Water Resources Research, vol. 33, n°12, p. 2909-2922.

- Merzouki, A., McNairn, H. et Pacheco, A. (2011) Mapping Soil Moisture Using RADARSAT-2 Data and Local Autocorrelation Statistics. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, n°1, p. 128-137.
- Moran, M. S., Peters-Lidard, C. D., Watts, J. M. et McElroy, S. (2004) Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian journal of remote sensing*, vol. 30, n°5, p. 805–826.
- Nelson, M. M. et Illingworth, W. T. (1991) *A practical guide to neural nets*. Reading, Mass : Addison-Wesley.
- NeuralWare, Inc. (1991) *Neural Computing, NeuralWorks Professional II/Plus and NeuralWorks Explorer*.
- Nguyen, D. et Widrow, B. (1990) Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Neural Networks, 1990., 1990 IJCNN International Joint Conference on*, p. 21-26 vol.3.
- Notarnicola, C., Angiulli, M. et Posa, F. (2008) Soil moisture retrieval from remotely sensed data: Neural network approach versus Bayesian method. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, n°2, p. 547-557.
- Oh, Y. (2004) Quantitative Retrieval of Soil Moisture Content and Surface Roughness From Multipolarized Radar Observations of Bare Soil Surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, n°3, p. 596-601.
- Oh, Y., Sarabandi, K. et Ulaby, F. T. (1992) An empirical model and an inversion technique for radar scattering from bare soil surfaces. *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 30, n°2, p. 370–381.
- Oh, Y., Sarabandi, K. et Ulaby, F. T. (1994) An inversion algorithm for retrieving soil moisture and surface roughness from polarimetric radar observation. *Geoscience and Remote Sensing Symposium, 1994. IGARSS'94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, vol. 3, p. 1582–1584. IEEE.
- Oh, Y., Sarabandi, K. et Ulaby, F. T. (2002) Semi-empirical model of the ensemble-averaged differential Mueller matrix for microwave backscattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, n°6, p. 1348-1355.
- Ontario Ministry of Natural Resources (2008) *Southern Ontario Land Resource Information System (SOLRIS) Phase 2 - Data Specifications, Version 1.2.*, 17 p.
- Paloscia, S. (2002) A summary of experimental results to assess the contribution of SAR for mapping vegetation biomass and soil moisture. *Canadian Journal of Remote Sensing*, vol. 28, n°2, p. 246–261.
- Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L. et Reppucci, A. (2013) Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. *Remote Sensing of Environment*, vol. 134, p. 234-248.
- Paris Anguela, T., Zribi, M., Hasenauer, S., Habets, F. et Loumagne, C. (2008) Analysis of surface and root-zone soil moisture dynamics with ERS scatterometer and the

- hydrometeorological model SAFRAN-ISBA-MODCOU at Grand Morin watershed (France). *Hydrology and Earth System Sciences*, vol. 12, n°6, p. 1415–1424.
- Parks Canada (2011) Trent-Severn Watersheds. Shoreline Policy and Regulation: Review and Recommendations (p. 121).
- Patel, P., Srivastava, H. S., Panigrahy, S. et Parihar, J. S. (2006) Comparative evaluation of the sensitivity of multi-polarized multi-frequency SAR backscatter to plant density. *International Journal of Remote Sensing*, vol. 27, n°2, p. 293–305.
- PCI Geomatics (2009) Geomatica Version 10.3 Users Manual. Ontario, Canada.
- Pierdicca, N., Castracane, P. et Pulvirenti, L. (2008) Inversion of Electromagnetic Models for Bare Soil Parameter Estimation from Multifrequency Polarimetric SAR Data. *Sensors*, vol. 8, n°12, p. 8181–8200.
- Planscape (2009) Holland Marsh agricultural impact study.
- Posa, F., Notarnicola, C. et Angiulli, M. (2004) Soil parameters retrieval from remotely sensed data: efficiency of neural network and Bayesian approaches. *Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International*, vol. 7, p. 4682–4685. IEEE.
- Quesney, A., Hégarat-Masclé, S. L., Taconet, O., Vidal-Madjar, D., Wigneron, J. P., Loumagne, C. et Normand, M. (2000) Estimation of Watershed Soil Moisture Index from ERS/SAR Data. *Remote Sensing of Environment*, vol. 72, n°3, p. 290–303.
- RADARSAT International (2000) RADARSAT Data Products Specifications, RSI-GS 026, Version 3/0.
- Rahman, M. M., Moran, M. S., Thoma, D. P., Bryant, R., Sano, E. E., Holifield Collins, C. D., Skirvin, S., Kershner, C. et Orr, B. J. (2007) A derivation of roughness correlation length for parameterizing radar backscatter models. *International Journal of Remote Sensing*, vol. 28, n°18, p. 3995–4012.
- Rodriguez-Fernandez, N. J. (2014) Communication personnelle. *International Geoscience and Remote Sensing Symposium*.
- Rodriguez-Fernandez, N. J., Aires, F., Richaume, P., Kerr, Y. H., Prigent, C., Kolassa, J., Cabot, F., Jimenez, C., Mahmoodi, A. et Drusch, M. (2015) Soil Moisture Retrieval Using Neural Networks: Application to SMOS. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, n°11, p. 5991–6007.
- Rojas, R. (1996) *Neural networks: a systematic introduction*. Berlin ; New York : Springer-Verlag. 502 p.
- Romshoo, S. A., Koike, M., Onaka, S., Oki, T. et Musiake, K. (2002) Influence of surface and vegetation characteristics on C-band radar measurements for soil moisture content. *Journal of the Indian Society of Remote Sensing*, vol. 30, n°4, p. 229–244.
- Said, S., Kothyari, U. et Arora, M. (2008) ANN-Based Soil Moisture Retrieval over Bare and Vegetated Areas Using ERS-2 SAR Data. *Journal of Hydrologic Engineering*, vol. 13, n°6, p. 461–475.

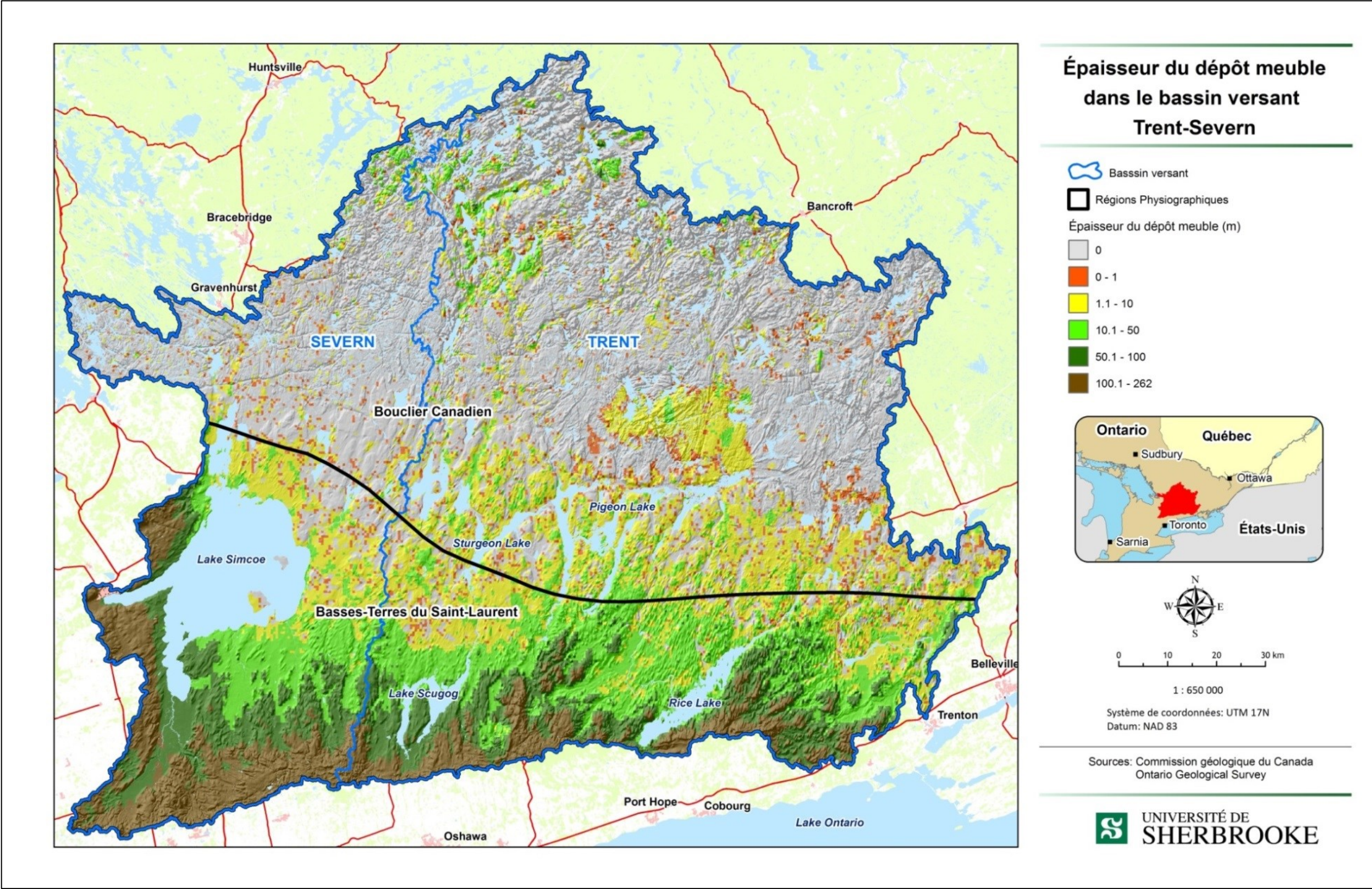
- Satalino, G., Mattia, F., Davidson, M. W. J., Thuy Le Toan, Pasquariello, G. et Borgeaud, M. (2002) On current limits of soil moisture retrieval from ERS-SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, n°11, p. 2438-2447.
- Schmugge, T. J. (1983) Remote sensing of soil moisture: Recent advances. *Geoscience and Remote Sensing, IEEE Transactions on*, n°3, p. 336-344.
- Sørensen, R., Zinko, U. et Seibert, J. (2006) On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences*, vol. 10, n°1, p. 101-112.
- Srivastava, H. S., Patel, P., Manchanda, M. L. et Adiga, S. (2003) Use of multiincidence angle RADARSAT-1 SAR data to incorporate the effect of surface roughness in soil moisture estimation. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, n°7, p. 1638-1640.
- Srivastava, H. S., Patel, P. et Naval Gund, R. R. (2006) Incorporating soil texture in soil moisture estimation from extended low-1 beam mode RADARSAT-1 SAR data. *International Journal of Remote Sensing*, vol. 27, n°12, p. 2587-2598.
- Swingler, K. (2001) *Applying neural networks: a practical guide* (3^e éd.). San Francisco : Kaufman.
- The MathWorks, Inc. (2014) *Neural network toolbox™ 6: User's guide* (R2014a).
- Thorstensen, A., Nguyen, P., Hsu, K. et Sorooshian, S. (2016) Using Densely Distributed Soil Moisture Observations for Calibration of a Hydrologic Model. *Journal of Hydrometeorology*, vol. 17, n°2, p. 571-590.
- Trudel, M., Leconte, R. et Charbonneau, F. (2012) Using RADARSAT-2 polarimetric and ENVISAT-ASAR dual-polarization data for estimating soil moisture over agricultural fields. *Canadian Journal of Remote Sensing*, vol. 38, n°4, p. 514-527.
- Ulaby, F. et Batlivala, P. (1976) Optimum Radar Parameters for Mapping Soil Moisture. *IEEE Transactions on Geoscience Electronics*, vol. 14, n°2, p. 81-93.
- Ulaby, F., Moore, R. et Fung, A. (1986) *Microwave remote sensing: Active and passive. Volume 3-From theory to applications*. Reading, MA : Addison-Wesley Publishing Company Inc.
- Ulaby, F. T. (1974) Radar measurement of soil moisture content. *Antennas and Propagation, IEEE Transactions on*, vol. 22, n°2, p. 257-265.
- Ulaby, F. T., Batlivala, P. P. et Dobson, M. C. (1978) Microwave backscatter dependence on surface roughness, soil moisture, and soil texture: Part I-bare soil. *IEEE Transactions on Geoscience Electronics*, vol. 16, n°4, p. 286-295.
- Ulaby, F. T., Bradley, G. A. et Dobson, M. C. (1979) Microwave Backscatter Dependence on Surface Roughness, Soil Moisture, and Soil Texture: Part II-Vegetation-Covered Soil. *IEEE Transactions on Geoscience Electronics*, vol. 17, n°2, p. 33-40.
- Ulaby, F. T., Dubois, P. C. et van Zyl, J. (1996) Radar mapping of surface soil moisture. *Journal of Hydrology*, vol. 184, n°1-2, p. 57-84.

- Ulaby, F. T., Moore, R. K. et Fung, A. K. (1982) Microwave remote sensing: Active and passive. Volume 2-Radar remote sensing and surface scattering and emission theory (Vol. 2). Reading, MA : Addison-Wesley Publishing Company Inc.
- Ulaby, F. T., Moore, R. K., Fung, A. K. et House, A. (1981) Microwave remote sensing: active and passive (Vol. 1-Microwave remote sensing fundamentals and radiometry). Reading, MA : Addison-Wesley Publishing Company Inc.
- van Zyl, J. (1989) Unsupervised classification of scattering behavior using radar polarimetry data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27, n°1, p. 36–45.
- Verhoest, N. E. ., Lievens, H., Wagner, W., Álvarez-Mozos, J., Moran, M. S. et Mattia, F. (2008) On the Soil Roughness Parameterization Problem in Soil Moisture Retrieval of Bare Surfaces from Synthetic Aperture Radar. *Sensors*, vol. 8, n°7, p. 4213-4248.
- Vina, A. (2004) Remote estimation of leaf area index and biomass in corn and soybean. Thèse de doctorat, University of Nebraska, Lincoln, États-Unis, 132 p.
- Vulfson, L., Genis, A., Blumberg, D. G., Sprintsin, M., Kotlyar, A., Freilikher, V. et Ben-Asher, J. (2012) Retrieval of surface roughness parameters of bare soil from the radar satellite data. *Journal of Arid Environments*, vol. 87, p. 77-84.
- Wagner, W., Lemoine, G. et Rott, H. (1999) A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. *Remote Sensing of Environment*, vol. 70, n°2, p. 191-207.
- Walker, J. P., Houser, P. R. et Willgoose, G. R. (2004) Active microwave remote sensing for soil moisture measurement: a field evaluation using ERS-2. *Hydrological Processes*, vol. 18, n°11, p. 1975-1997.
- Wang, J. R. et Choudhury, B. J. (1995) Passive microwave radiation from soil: Examples of emission model and observations,. Zeist, The Netherlands : VSP.
- Wang, J. R., Hsu, A., Shi, J. C., O'Neill, P. E. et Engman, E. T. (1997) A comparison of soil moisture retrieval models using SIR-C measurements over the little Washita River watershed. *Remote Sensing of Environment*, vol. 59, n°2, p. 308-320.
- Weiss, M. (1998) Development of an algorithm for large scale vegetation monitoring (Theses). Thèse de doctorat, Université Nice Sophia Antipolis, 188 p.
- Wilson, D. J., Western, A. W., Grayson, R. B., Berg, A. A., Lear, M. S., Rodell, M., Famiglietti, J. S., Woods, R. A. et McMahon, T. A. (2003) Spatial distribution of soil moisture over 6 and 30cm depth, Mahurangi river catchment, New Zealand. *Journal of Hydrology*, vol. 276, n°1-4, p. 254-274.
- Yamaguchi, Y., Moriyama, T., Ishido, M. et Yamada, H. (2005) Four-component scattering model for polarimetric SAR image decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, n°8, p. 1699-1706.
- Zehe, E. et Blöschl, G. (2004) Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions: PREDICTABILITY OF HYDROLOGIC RESPONSE. *Water Resources Research*, vol. 40, n°10, p. n/a-n/a.

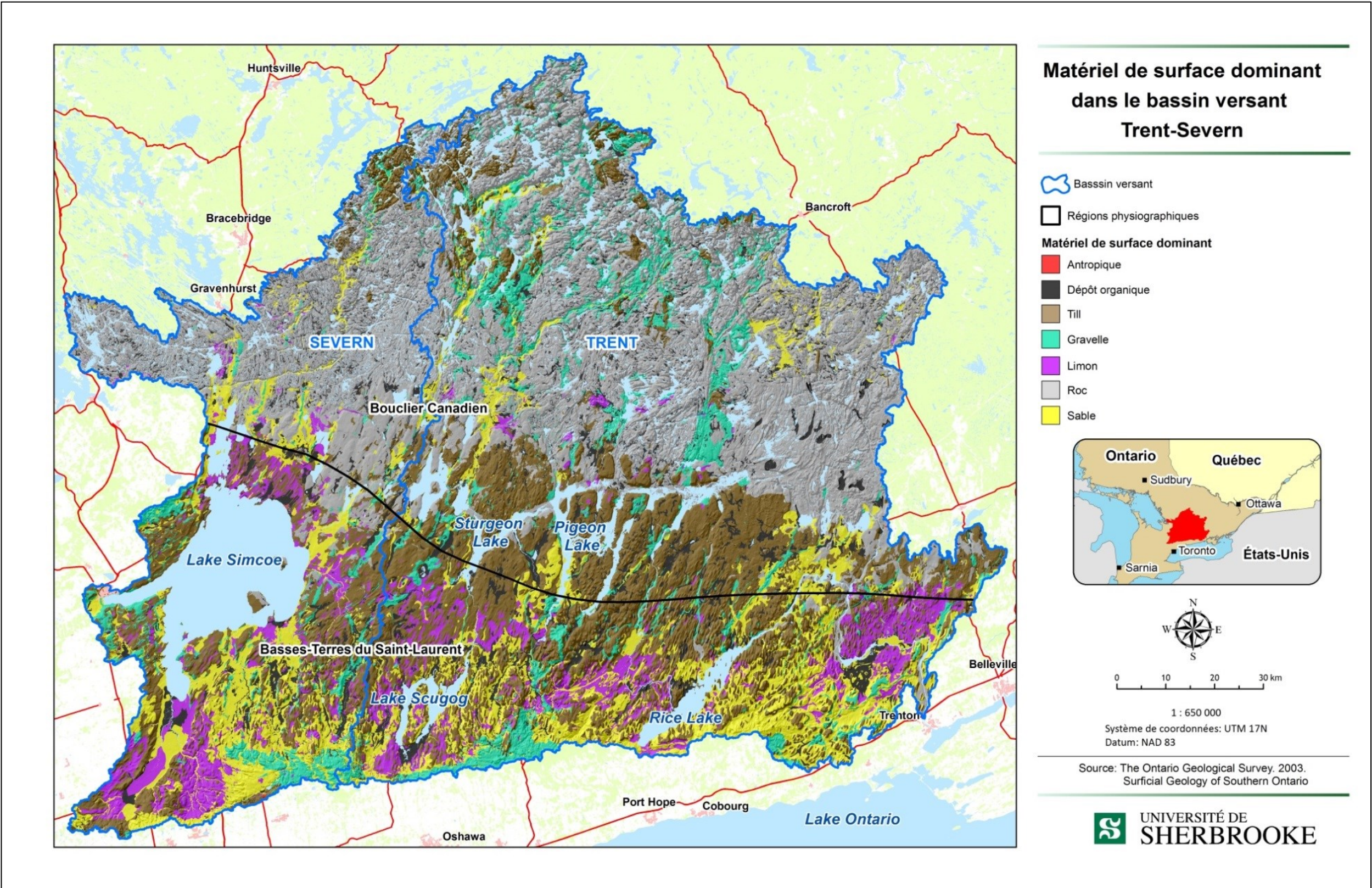
Zribi, M. et Dechambre, M. (2002) A new empirical model to retrieve soil moisture and roughness from C-band radar data. *Remote Sensing of Environment*, vol. 84, n°1, p. 42–52.

ANNEXE A. CARTOGRAPHIE DE LA ZONE D'ÉTUDE

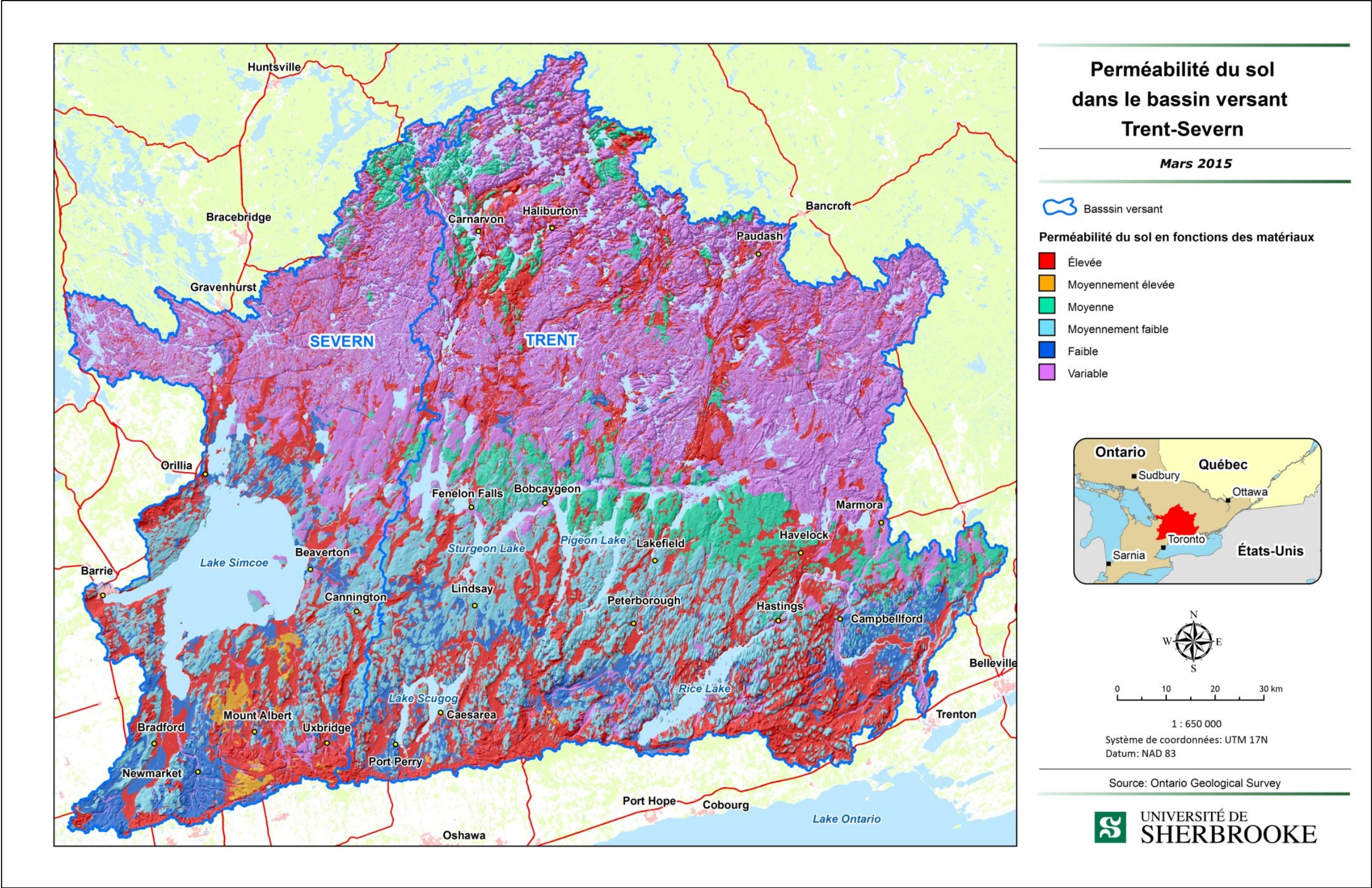
A.1 ÉPAISSEUR DU DÉPÔT MEUBLE



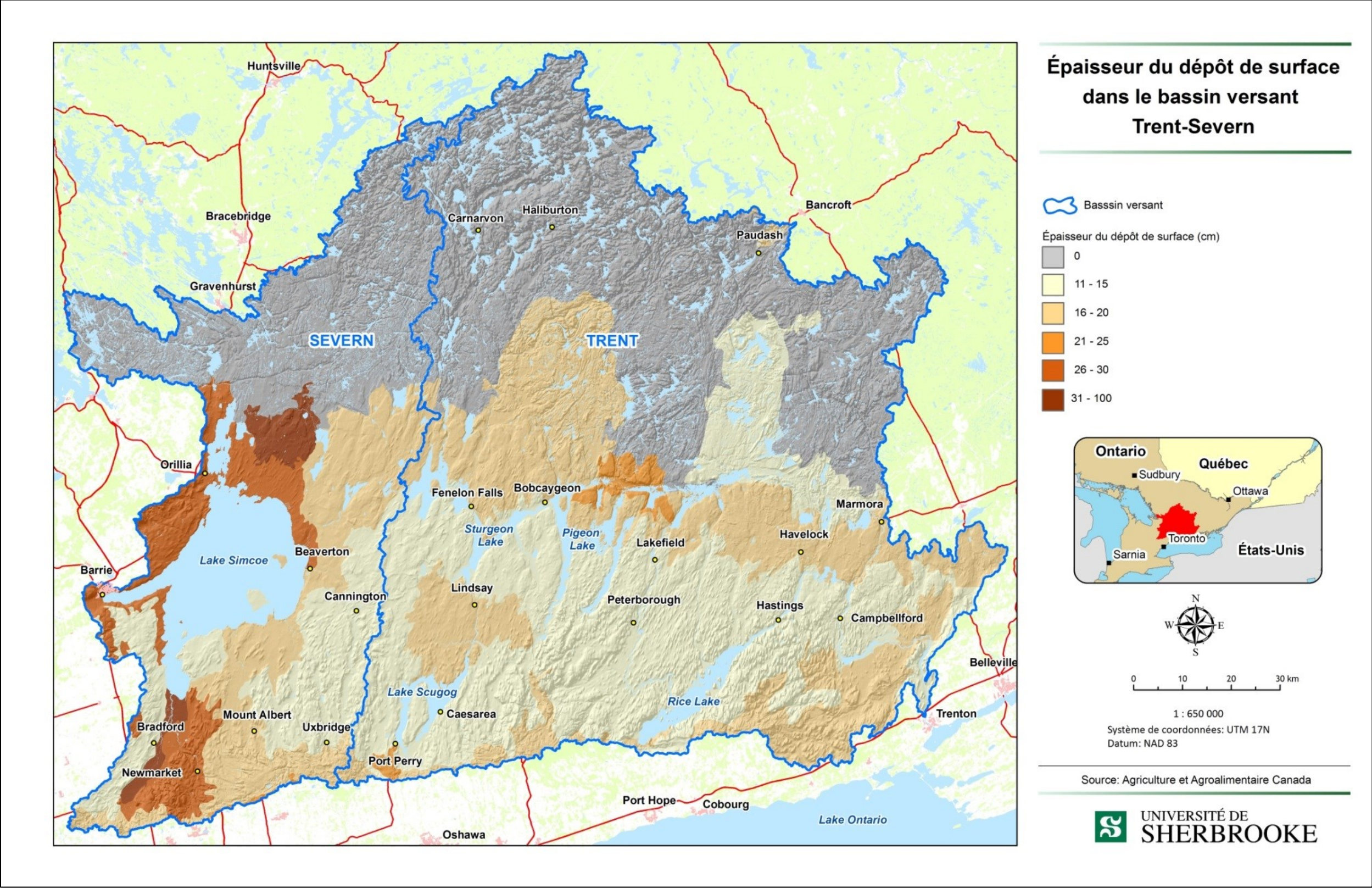
A.2 MATÉRIEL DE SURFACE DOMINANT



A.3 PERMÉABILITÉ DU SOL

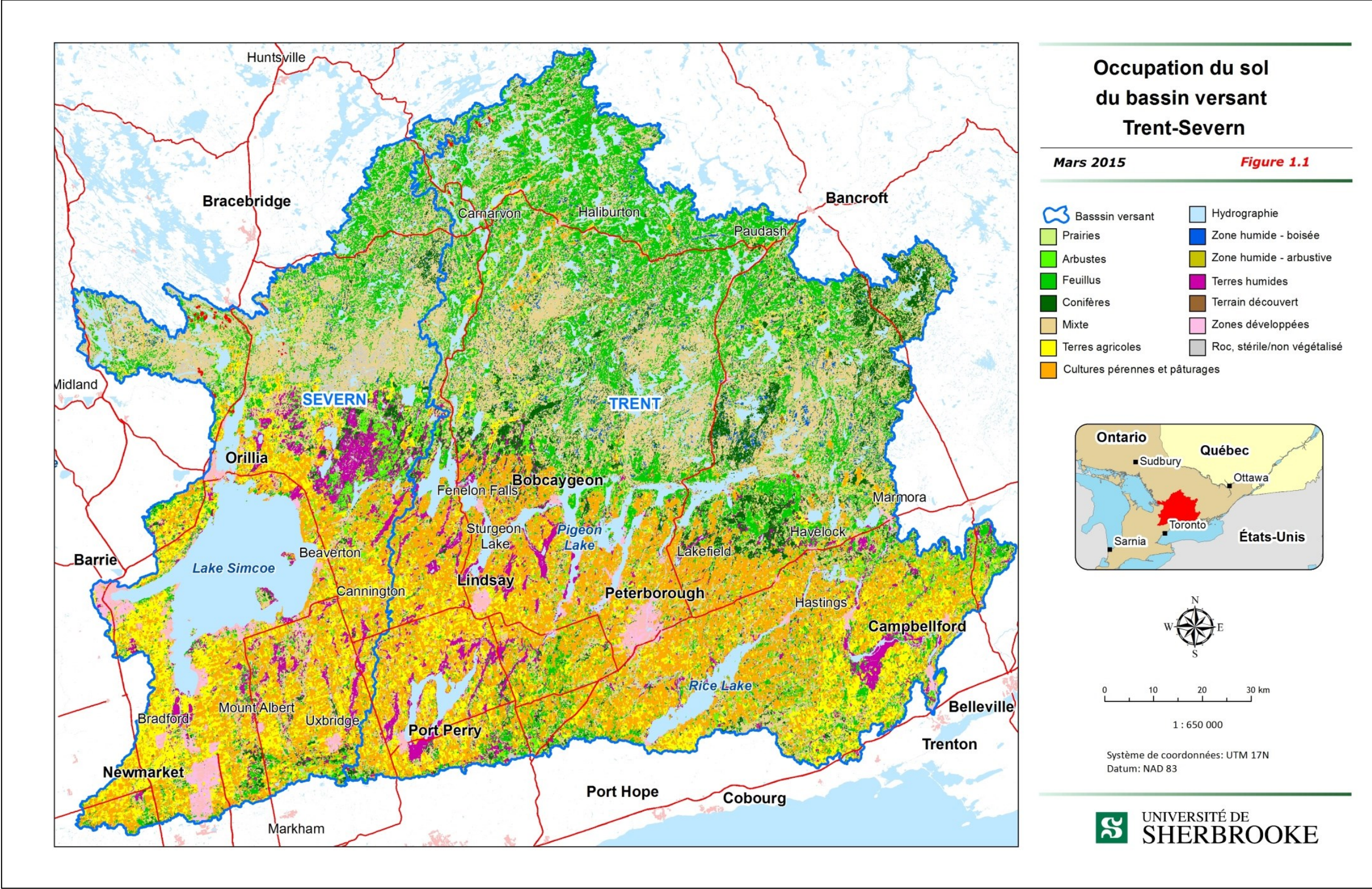


A.4 ÉPAISSEUR DU DÉPÔT DE SURFACE

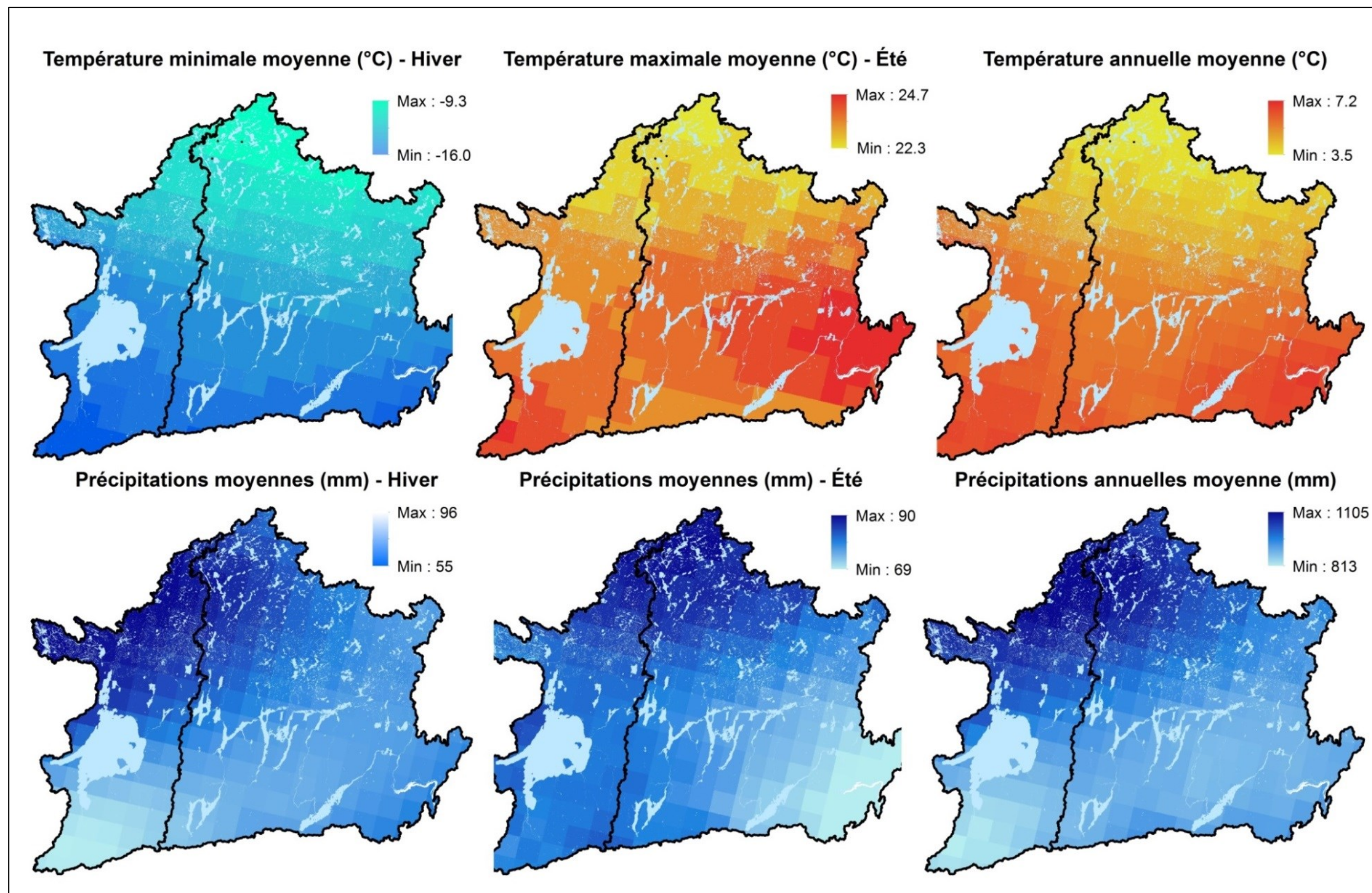




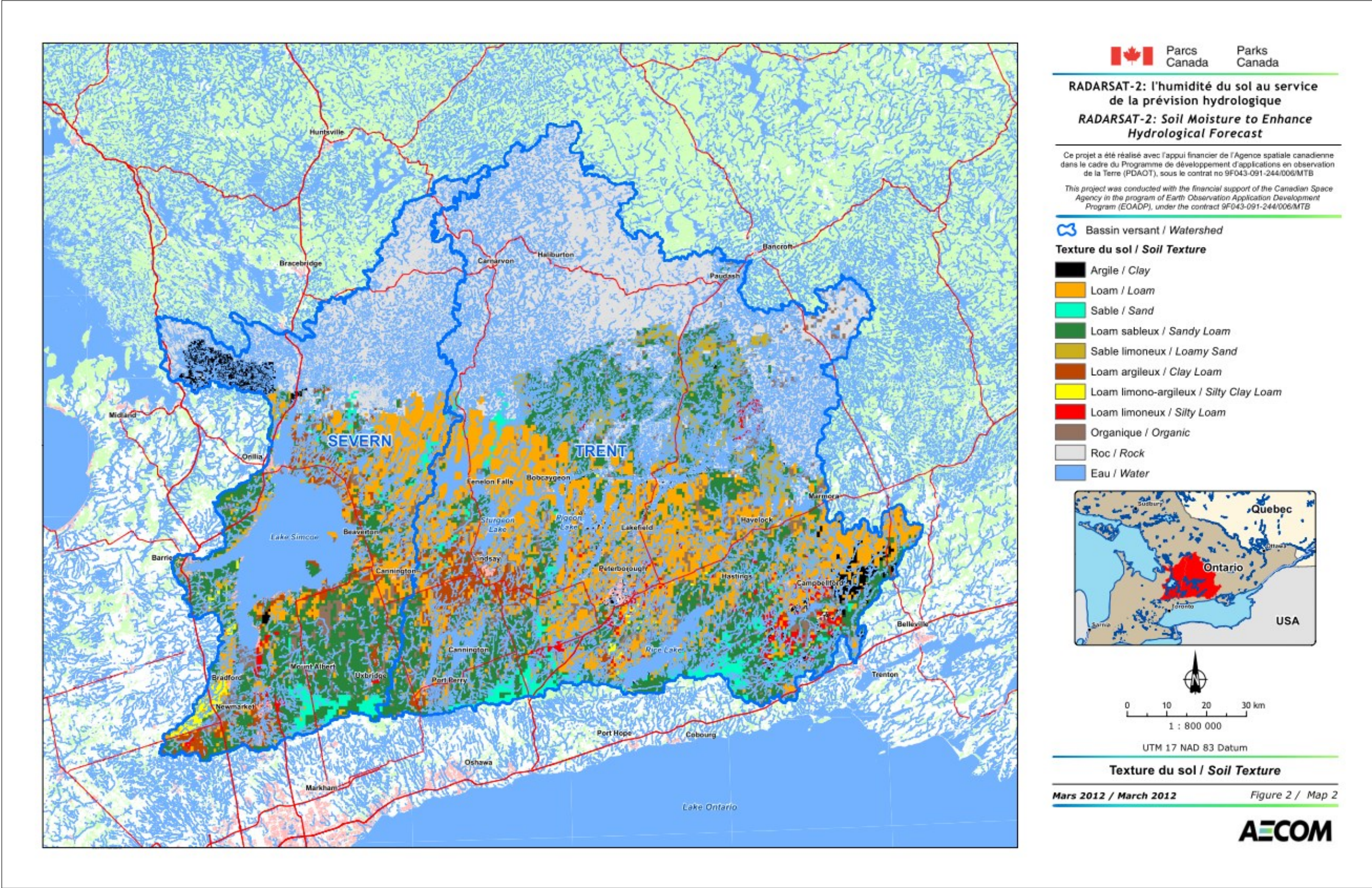
A.6 OCCUPATION DU SOL



A.7 CLIMAT

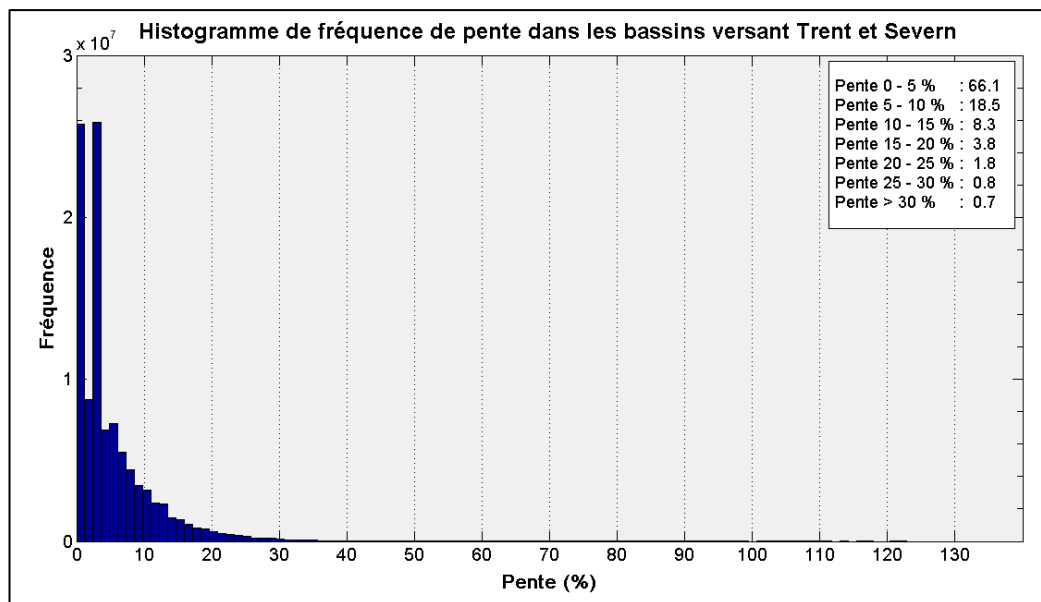


A.8 TEXTURE DU SOL



ANNEXE B. HISTOGRAMME DE FRÉQUENCE DE PENTE

DE PENTE



ANNEXE C. SPÉCIFICATIONS DES MNÉ

Jeu de donnée	Précision (m)		Note sur la précision verticale	Altitude (m)	
	Horizontale	Verticale		Min	Max
030M13	5	3	*	132	481
030M14	1	1	*	73	355
031C04	5	3	*	71	311
031C05	5	3	*	101	286
031C06	5	3	*	78	221
031C12	5	3	*	171	381
031C13	5	3	*	273	450
031C14	5	3	*	233	450
031D01	5	3	*	108	357
031D02	5	3	*	164	396
031D03	5	3	*	217	410
031D04	5	3	*	198	385
031D05	5	3	*	174	344
031D06	5	3	*	217	324
031D07	5	3	*	240	361
031D08	5	3	*	184	328
031D09	5	3	*	230	376
031D10	5	3	*	214	342
031D11	5	3	*	184	317

Jeu de donnée	Précision (m)		Note sur la précision verticale	Altitude (m)	
	Horizontale	Verticale		Min	Max
031D12	5	3	*	171	411
031D13	5	3	*	174	326
031D14	5	3	*	202	346
031D15	5	3	*	255	410
031D16	5	3	*	292	512
031E01	5	3	*	326	541
031E02	5	3	*	290	496
031E03	5	3	*	221	407
031E07	5	3	*	307	562
031E08	5	3	*	322	553
031F04	5	3	*	276	518

*inconnue pour certaines sources

en général, 3 pour les courbes et 1 pour les points cotés